

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM Protein - Protein search, using sw model

Run on: July 9, 2004, 12:59:40 ; Search time 55 Seconds  
(without alignments)  
570,232 Million cell updates/sec

Title: US-09-784-553C-19  
Perfect score: 605

Sequence: 1 KFGRVITNQLOYLHKVVMKAL.....IVLMACTLERIFLQKVASMP 111

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

Database : A\_Geneseq\_29Jan04:  
1: \_GeneseqGP1805:/\*  
2: \_GeneseqGP1990s:/\*  
3: \_GeneseqGP2000s:/\*  
4: \_GeneseqGP2001s:/\*  
5: \_GeneseqGP2002s:/\*  
6: \_GeneseqGP2003as:/\*  
7: \_GeneseqGP2003bs:/\*  
8: \_GeneseqGP2004s:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	754	2 AAY07027	Aay07027 Breast, Ca
2	598	100.0	801	7 ADC35076	Adc35076 Human bre
3	598	98.8	303	4 AAB58502	Aab58502 Human pro
4	542	89.6	140	4 AAU16619	Aau16619 Human nov
5	542	89.6	140	6 ABUS5688	Abu5688 Human nov
6	542	89.6	235	4 AAU16206	Aau16206 Human nov
7	542	89.6	235	6 ABUS5275	Abu5275 Human nov
8	500	82.6	688	3 AAY157898	Aay157898 Human tra
9	500	82.6	1390	5 ABU65152	Abu65152 Human NOV
10	498	82.3	1937	4 AB588985	Abb588985 Drosophil
11	459	75.9	947	2 AAN81168	Aaw81168 Transcrip
12	459	75.9	947	2 AAY07114	Aay07114 WO904265
13	281	62.6	573	6 ABP41326	Abp41326 Human DIT
14	281	46.5	513	4 ABB68341	Abb68341 Drosophil
15	269	44.5	330	6 ABR1589	Abri1589 Human DIT
16	237	39.2	648	4 ABB53109	Abb53109 Drosophil
17	221	36.5	703	5 ABB93466	Abb93466 Herbicida
18	208	34.4	743	3 AAG32556	Aag32556 Arabidops
19	208	34.4	764	3 AAG52555	Aag52555 Arabidops
20	208	34.4	813	3 AAG52554	Aag52554 Arabidops
21	207	34.2	124	3 AAG12245	Aag12245 Arabidops
22	204	33.7	270	3 AAG42244	Aag42244 Arabidops
23	204	33.7	386	3 AAG42243	Aag42243 Arabidops
24	204	33.7	386	6 ABBP70522	Abp70522 Amino aci
25	33.7				

#### RESULT 1

ID AA107027 standard; protein; 754 AA.

XX AAY07027;

XX DT 02-JUL-1999 (first entry)

XX DE Breast cancer associated antigen precursor sequence.  
XX Cancer associated antigen; diagnosis; research; treatment; human;  
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
KW prostate cancer.  
XX Homo sapiens.  
XX OS WO9904265-A2.  
XX PN WO9904265-A2.

#### RESULT 2

ID AA107027 standard; protein; 754 AA.

XX AC AAY07027;

XX DT 02-JUL-1999 (first entry)

XX DE Breast cancer associated antigen precursor sequence.  
XX Cancer associated antigen; diagnosis; research; treatment; human;  
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;  
KW prostate cancer.  
XX Homo sapiens.  
XX OS WO9904265-A2.  
XX PN WO9904265-A2.

#### ALIGNMENTS

Abb33927 Herbicida  
Abb3654 Eucalyptu  
Aaw40054 P300/CBP-  
Aau16618 Human nov  
Abu55687 Human nov  
Aaw40052 Human P30  
Aaw40052 SCA/N/CB8  
Aau73845 SCA/N/CB8  
Aau13446 Human dia  
Aay04123 Human sec  
Aao16118 Human nuc  
Aay57453 Human tra  
Aay57452 Human tra  
Aag36672 Arabidops  
Aag36671 Arabidops  
Aag41764 Arabidops  
Aag41763 Arabidops  
Aag41762 Arabidops  
Aag36673 Arabidops  
Aar79054 CB8 bind  
Aaw40058 Cellular

The invention relates to a method for diagnosing a disorder characterised by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a nucleic acid sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the

CC

CC interaction between the agent and the NAM or the expression product as a  
 CC determination of the disorder. The products and methods can be used in  
 CC the diagnosis, monitoring, research, or treatment of conditions  
 CC characterised by the expression of various cancer associated antigens.  
 CC The invention provides nucleic acid sequences and encoded polypeptides  
 CC which are cancer associated antigen precursors expressed in human breast  
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and  
 CC lung cancer  
 XX Sequence 754 AA;

Query Match Similarity 100.0%; Score 605; DB 2; Length 754;

Best Local Similarity 100.0%; Pred. No. 2.8e-68; Mismatches 0; Gaps 0;

Matches 111; Conservative 0; N mismatches 0; Indels 0; Gaps 0;

Qy 1 KPGRVINQOLYIHKYVMKALWKHQFAWPFRQPAVKGGLPDYHKIKQPMGNTIKRRL 60

Db 24 KPGRVINQOLYIHKYVMKALWKHQFAWPFRQPAVKGGLPDYHKIKQPMGNTIKRRL 83

Qy 61 ENNYWAASECMQDENTMFTNCYLYNKPTDDIVMAQTLKEFLQYKASMP 111

Db 84 ENNYWAASECMQDENTMFTNCYLYNKPTDDIVMAQTLKEFLQYKASMP 134

RESULT 2  
 ADC35076 standard; protein: 801 AA.  
 ID ADC35076  
 XX DT 18-DEC-2003 (first entry)  
 AC ADC35076  
 XX DE Human breast cancer antigen seq id 42.  
 XX PR 11-JAN-2000; 20000JP-00118767;  
 KW breast cancer; breast cancer diagnosis; breast cancer antigen.  
 XX PR 02-MAY-2000; 20000JP-00183757;  
 OS Homo sapiens.  
 XX PR 09-JUN-2000; 20000JP-00241899.  
 PA (HELI-) HELIX RES INST.  
 XX PR Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PR Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX DR 20000JP-00116126.  
 XX PR 29-JUL-1999; 990JP-00248036.  
 PR 27-AUG-1999; 990JP-00300253.  
 PR 11-JAN-2000; 20000JP-00118767;  
 PR 02-MAY-2000; 20000JP-00183757;  
 PR 09-JUN-2000; 20000JP-00241899.  
 PA (HELI-) HELIX RES INST.

XX PR 12-JUN-2003.  
 PD 15-MAY-2002; 20020US-00146473.  
 XX PR 15-MAY-2001; 20010US-0291150P.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX PI Scanian MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;  
 DR 2003-829397/77.  
 XX PR N-PSDB; ADC35118.

PT Diagnosing breast cancer in subject by obtaining biological sample from  
 PT subject, contacting sample with breast cancer-associated polypeptides,  
 PT determining specific binding between polypeptides and agents in sample.  
 PA Example 2: SEQ ID NO 42; 173pp; English.  
 XX PI The invention describes a method of diagnosing breast cancer in subject  
 CC comprising contacting biological sample from subject with at least two  
 CC different breast cancer-associated polypeptides (1) encoded by nucleic  
 CC acid molecules (II) comprising sequence chosen from sequence 42 fully defined  
 CC sequences as given in specification, determining specific binding between  
 CC (1) and agents in sample, where presence of the binding is diagnostic for  
 CC breast cancer. The method is useful for diagnosing breast cancer in a  
 CC subject. The sample is blood, lymph node fluid or breast discharge fluid.  
 CC This is the amino acid sequence of a breast cancer antigen.  
 XX SQ Sequence 801 AA;

Query Match Similarity 100.0%; Score 605; DB 7; Length 801;

Best Local Similarity 100.0%; Pred. No. 3.1e-68;

CC Matches 111; Conservative 0; N mismatches 0; Indels 0; Gaps 0;  
 CC Qy 1 KPGRVINQOLYIHKYVMKALWKHQFAWPFRQPAVKGGLPDYHKIKQPMGNTIKRRL 60  
 CC DB 71 KPGRVINQOLYIHKYVMKALWKHQFAWPFRQPAVKGGLPDYHKIKQPMGNTIKRRL 130  
 CC 61 ENNYWAASECMQDENTMFTNCYLYNKPTDDIVMAQTLKEFLQYKASMP 111  
 CC DB 131 ENNYWAASECMQDENTMFTNCYLYNKPTDDIVMAQTLKEFLQYKASMP 181  
 CC  
 RESULT 3  
 AAB95802 standard; protein: 303 AA.  
 ID AAB95802  
 XX AC AAB95802;  
 XX DT 26-JUN-2001 (first entry)  
 XX DE Human protein sequence SEQ ID NO:18783.  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX OS Homo sapiens.  
 XX PR EP1074617-A2.  
 XX PD 07-FEB-2001.  
 XX PR 28-JUL-2000; 20000EP-00116126.  
 XX PR 29-JUL-1999; 990JP-00248036.  
 PR 27-AUG-1999; 990JP-00300253.  
 PR 11-JAN-2000; 20000JP-00118767;  
 PR 02-MAY-2000; 20000JP-00183757;  
 PR 09-JUN-2000; 20000JP-00241899.  
 PA (HELI-) HELIX RES INST.  
 XX PR Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PR Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX DR 20000JP-00116126.  
 XX PR 29-JUL-1999; 990JP-00248036.  
 PR 27-AUG-1999; 990JP-00300253.  
 PR 11-JAN-2000; 20000JP-00118767;  
 PR 02-MAY-2000; 20000JP-00183757;  
 PR 09-JUN-2000; 20000JP-00241899.  
 PA (HELI-) HELIX RES INST.

Claim 8; SEQ ID NO 18783; 2537pp + Sequence Listing; English.  
 XX DR 2001-318749/34.  
 XX PR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
 XX PT The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the 5'-end sequence of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13678 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH332 represent human amino acid sequences; and AAH3629 to AAH332 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 303 AA;  
 & Query Match 98.8%; Score 598; DB 4; Length 303;  
 Best Local Similarity 99.1%; Pred. No. 6.7e-68;  
 Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 110;  
 Matches 110;  
 QY 1 KPGRTVNLQYIHKVYKALWKHQPAWPFRPVDAYKLGLPDYHKKIKQFMDGTTKERRQL 60  
 Db 74 KPGRTVNLQYIHKVYKALWKHQPAWPFRPVDAYKLGLPDYHKKIKQFMDGTTKERRQL 133  
 QY 61 ENYYXWAASECMQDENTMFNTCYIYNKPTTDIVMAQTPEKIFLOKVASMP 111  
 Db 134 ENYYXWAASECMQDENTMFNTCYIYNKPTTDIVMAQTPEKIFLOKVASMP 184

RESULT 4  
 AAU16619 standard; protein: 140 AA.  
 AAU16619;  
 XX DT 07-NOV-2001 (first entry)  
 XX Human novel secreted protein, Seq ID 1572.  
 DE Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
 KW cardiotonic; vasotropic; cerebroprotective; nootropic; neuroprotective;  
 KW antibacterial; virucide; fungicide; opthalmological; vulnerary;  
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW cerebroischaemia; angiogenesis; nervous system disorder;  
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
 KW preservative; antiproliferative.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX WO20015322-A2.  
 XX 02-AUG-2001.  
 XX 17-JAN-2001; 2001WO-US001341.  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-MAR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0221518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226219P.  
 PR 22-AUG-2000; 2000US-0226651P.  
 PR 22-AUG-2000; 2000US-022688P.  
 PR 23-AUG-2000; 2000US-0227099P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229813P.  
 PR 01-SEP-2000; 2000US-0229333P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230432P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232030P.  
 PR 08-SEP-2000; 2000US-0232031P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-02363802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 13-OCT-2000; 2000US-0239337P.  
 PR 20-OCT-2000; 2000US-024160P.  
 PR 20-OCT-2000; 2000US-024121P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 08-NOV-2000; 2000US-0241826P.  
 PR 08-NOV-2000; 2000US-024417P.  
 PR 08-NOV-2000; 2000US-0244674P.  
 PR 08-NOV-2000; 2000US-0244675P.  
 PR 08-NOV-2000; 2000US-024525P.  
 PR 08-NOV-2000; 2000US-024526P.  
 PR 08-NOV-2000; 2000US-024617P.  
 PR 08-NOV-2000; 2000US-024617P.  
 PR 08-NOV-2000; 2000US-0246578P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246513P.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0250392P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251189P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX PA (HUNYA-) HUMAN GENOME SCI INC.  
 XX P1 Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-489783/53.  
 DR N-PSDB; AAS26606.  
 XX PT New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.  
 XX P1 SEQ ID NO 1572; 980pp; English.  
 XX CC The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms of the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides, can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative. The polypeptides can also be used to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, co-factors and other nutritional components. The present sequence represents a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 89.6%; Score 542; DB 4; Length 140;  
 Best Local Similarity 87.4%; Pred. No. 3.7e-61;  
 Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRVITNQLOQYLHKYTMKALMKHQFAWPFRQPVDAVKGKLPDYHKIKPQMDMGTIKRRL 60  
 DB 5 KERGRKTNHQLOQMNYVVTKLWKRHQFAWPFYQPVDAIKLNPDYHKIKNPMQMDMGTIKRRL 64

QY 61 ENYYYNASECMQDENTMFPINCYTINKPTDIVMAQTEKIFLOKVASMP 111  
 DB 65 ENYYYNASECMQDENTMFPINCYTINKPTDIVMAQLEKIFLOKVAQMP 115

RESULT 5  
 ABU55688  
 ID ABU55688 standard; protein; 140 AA.  
 XX  
 AC ABU55688;  
 XX  
 AC Human novel polypeptide #775.  
 XX  
 DE Human novel polypeptide #775.  
 XX  
 KW Human, neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reactivity;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemotatic; antiarteriosclerotic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002132753-A1.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PR 17-JAN-2001; 2001US-0074864.  
 XX  
 PR 31-JAN-2000; 2000US-017905P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 07-JUL-2000; 2000US-0214647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0217496P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 11-AUG-2000; 2000US-0221480P.  
 PR 14-AUG-2000; 2000US-0221482P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225368P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225575P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0234273P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.

PR	29-SEP-2000;	2000US-0236369P.	XX	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic; cardiotropic; vasotropic; cereoprotective; nootropic; neuroprotective; antibacterial; virucidal; fungicidal; opthalmological; pulmonary; secreted protein; rheumatoid arthritis; hyperproliferative disorder; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder; corneal infection; wound healing; epithelial cell proliferation; skin ageing; food additive; preservative; anti-proliferative.
PR	02-OCT-2000;	2000US-0236810P.	XX	
PR	02-OCT-2000;	2000US-0237037P.	XX	
PR	02-OCT-2000;	2000US-0237038P.	XX	
PR	02-OCT-2000;	2000US-0237039P.	XX	
PR	02-OCT-2000;	2000US-0237040P.	XX	
PR	13-OCT-2000;	2000US-0239935P.	XX	
PR	20-OCT-2000;	2000US-0240960P.	XX	
PR	20-OCT-2000;	2000US-0241785P.	XX	
PR	01-NOV-2000;	2000US-0241809P.	XX	
PR	17-NOV-2000;	2000US-0244617P.	XX	
PR	08-DEC-2000;	2000US-0249239P.	XX	
PR	08-DEC-2000;	2000US-0251856P.	XX	
PR	08-DEC-2000;	2000US-0251868P.	XX	
PR	08-DEC-2000;	2000US-0251869P.	XX	
PA	(ROSE/)	ROSEN C A.	XX	17-JAN-2001; 2001WO-US001341.
PA	(RUBE/)	RUBEN S M.	XX	
PA	(BARA/)	BARASH S C.	XX	
PT	Rosen CA,	Ruben SM,	Barash SC;	
XX	WPI;	2003-147444/14.	XX	
DR	N-PSDB;	ABX73947.	XX	
XX	New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or renal disorders.	PT	PR	31-JAN-2000; 2000US-0179065P.
PT	PT	PR	04-FEB-2000; 2000US-0180628P.	
PT	PT	PR	24-FEB-2000; 2000US-0184664P.	
PT	PT	PR	02-MAR-2000; 2000US-0186350P.	
PT	PT	PR	16-MAR-2000; 2000US-018974P.	
PT	PT	PR	17-MAR-2000; 2000US-019076P.	
PT	PT	PR	18-APR-2000; 2000US-0198123P.	
PT	PT	PR	19-MAY-2000; 2000US-0201123P.	
PT	PT	PR	07-JUN-2000; 2000US-0205515P.	
PT	PT	PR	07-JUN-2000; 2000US-0209467P.	
PT	PT	PR	28-JUN-2000; 2000US-0214866P.	
PT	PT	PR	30-JUN-2000; 2000US-0215135P.	
PT	PT	PR	07-JUL-2000; 2000US-021647P.	
PT	PT	PR	07-JUL-2000; 2000US-021680P.	
PT	PT	PR	11-JUL-2000; 2000US-0217487P.	
PT	PT	PR	11-JUL-2000; 2000US-0217496P.	
CC	CC	PR	14-JUL-2000; 2000US-0218290P.	
CC	CC	PR	26-JUL-2000; 2000US-0220963P.	
CC	CC	PR	26-JUL-2000; 2000US-0220964P.	
CC	CC	PR	14-AUG-2000; 2000US-0224518P.	
CC	CC	PR	14-AUG-2000; 2000US-0224519P.	
CC	CC	PR	14-AUG-2000; 2000US-0225213P.	
CC	CC	PR	14-AUG-2000; 2000US-0225214P.	
CC	CC	PR	14-AUG-2000; 2000US-0225226P.	
CC	CC	PR	14-AUG-2000; 2000US-0225267P.	
CC	CC	PR	14-AUG-2000; 2000US-0225268P.	
CC	CC	PR	14-AUG-2000; 2000US-0225270P.	
CC	CC	PR	14-AUG-2000; 2000US-0225447P.	
CC	CC	PR	14-AUG-2000; 2000US-0225575P.	
CC	CC	PR	14-AUG-2000; 2000US-0225758P.	
CC	CC	PR	14-AUG-2000; 2000US-0225759P.	
CC	CC	PR	18-AUG-2000; 2000US-02256279P.	
CC	CC	PR	22-AUG-2000; 2000US-0225681P.	
CC	CC	PR	22-AUG-2000; 2000US-0227182P.	
CC	CC	PR	23-AUG-2000; 2000US-0227109P.	
CC	CC	PR	30-AUG-2000; 2000US-0228924P.	
QY	1	KPGRVTVNQIQLYIHKVNMKALMKHQFANPDPYDAVKIGLPDYHKIKLQPKMDMGTIKKR 60	PR	01-SEP-2000; 2000US-0229287P.
Db	5	KPGRKTNQIQLYQMVQMVVQVTLWQFANPDPYQVDAIKLNLPDYHKIKLQPKMDMGTIKKR 64	PR	01-SEP-2000; 2000US-0229343P.
Qy	61	ENNYWAASECMDENTNETNCYIYNKTDIVMAQTLKEFLQKVAQMP 111	PR	01-SEP-2000; 2000US-0229545P.
Db	65	ENNYWAASECMDENTNETNCYIYNKTDIVMAQALEKTFLQKVAQMP 115	PR	08-SEP-2000; 2000US-0230438P.
Qy	66	ENNYWAASECMDENTNETNCYIYNKTDIVMAQALEKTFLQKVAQMP 116	PR	08-SEP-2000; 2000US-0231244P.
Db	66	ENNYWAASECMDENTNETNCYIYNKTDIVMAQALEKTFLQKVAQMP 117	PR	08-SEP-2000; 2000US-0231413P.
AC	AAU16206;	AAU16206 standard; protein; 235 AA.	XX	08-SEP-2000; 2000US-0231414P.
AC	AAU16206;	AAU16206;	XX	08-SEP-2000; 2000US-0232080P.
AC	AAU16206;	AAU16206;	XX	08-SEP-2000; 2000US-0232081P.
DT	07-NOV-2001	(first entry)	XX	12-SEP-2000; 2000US-0232397P.
DE	Human novel secreted protein, Seq ID 1159.	XX	14-SEP-2000; 2000US-0232398P.	

RESULT 6

AAU16206  
AAU16206 standard; protein; 235 AA.

ID AAU16206

07-NOV-2001

(first entry)

XX Human novel secreted protein, Seq ID 1159.

14-SEP-2000; 2000US-0232399P.  
 PR PR 05-DEC-2000; 2000US-0256719P.  
 PR PR 06-DEC-2000; 2000US-025149P.  
 PR PR 08-DEC-2000; 2000US-0251856P.  
 PR PR 08-DEC-2000; 2000US-0251868P.  
 PR PR 08-DEC-2000; 2000US-025189P.  
 PR PR 08-DEC-2000; 2000US-025198P.  
 PR PR 08-DEC-2000; 2000US-0251990P.  
 PR PR 11-DEC-2000; 2000US-0254097P.  
 PR PR 05-JAN-2001; 2001US-0259678P.  
 XX XX (HUMA ) HUMAN GENOME SCI INC.  
 PA PA  
 PI PI Rosen CA, Barash SC, Ruben SM;  
 XX XX  
 WPI; 2001-488783/53.  
 DR DR N-PSDD; AAS26133.  
 XX XX  
 PR New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives.  
 XX XX  
 PS Claim 11; SEQ ID NO 1159; 980pp; English.  
 XX XX  
 CC The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Antibodies to the proteins can also be used in  
 CC alleviating symptoms associated with the disorders and in diagnostic  
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
 CC and many other disorders listed in the specification. The polypeptides  
 CC can also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence represents a novel secreted protein of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC  
 Query Match 99.6%; Score 542; DB 4; Length 235;  
 Best Local Similarity 87.4%; Pred. No. 7.4e-61;  
 Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Oy 1 KPGRTVNQOYLHRVVMKHOFAWPPRQPDAVKGLPDYHKITKOPMDNGTIRKRL 60  
 Db 5 KPGRTVNQOYOMVNTVVKTMWQFAMFYQPVDAIKUNLPDVKIKNPMDGTTIRKRL 64  
 61 ENNYIWSASECMQDFNTMFTNCIYINKPDTDIVMAQTEKIPFLQKVYASMP 111  
 65 ENNYIWSASECMQDFNTMFTNCIYINKPDTDIVMAQALEKIFLQKVYAQMP 115

RESULT 7  
 ABU55275 standard; protein; 235 AA.  
 ID ABU55275  
 XX  
 AC ABU55275;  
 XX  
 DT 18-MAR-2003 (first entry)  
 DE  
 XX  
 Human novel polypeptide #362.  
 XX  
 Human; neural disorder; immune system disorder; renal disorder;  
 muscular disorder; respiratory disease; reproductive disorder;  
 KW

PR 14-SEP-2000; 2000US-0232400P.  
 PR PR 14-SEP-2000; 2000US-0232401P.  
 PR PR 14-SEP-2000; 2000US-0233063P.  
 PR PR 14-SEP-2000; 2000US-0233064P.  
 PR PR 14-SEP-2000; 2000US-0233065P.  
 PR PR 21-SEP-2000; 2000US-0234223P.  
 PR PR 21-SEP-2000; 2000US-0234474P.  
 PR PR 25-SEP-2000; 2000US-0234997P.  
 PR PR 26-SEP-2000; 2000US-0234998P.  
 PR PR 27-SEP-2000; 2000US-0234984P.  
 PR PR 27-SEP-2000; 2000US-0235834P.  
 PR PR 29-SEP-2000; 2000US-0235836P.  
 PR PR 29-SEP-2000; 2000US-0236168P.  
 PR PR 29-SEP-2000; 2000US-0236336P.  
 PR PR 02-OCT-2000; 2000US-0236370P.  
 PR PR 02-OCT-2000; 2000US-0236802P.  
 PR PR 02-OCT-2000; 2000US-0237037P.  
 PR PR 02-OCT-2000; 2000US-0237040P.  
 PR PR 02-OCT-2000; 2000US-0237039P.  
 PR PR 13-OCT-2000; 2000US-0239337P.  
 PR PR 20-OCT-2000; 2000US-0240160P.  
 PR PR 20-OCT-2000; 2000US-0241221P.  
 PR PR 20-OCT-2000; 2000US-0241785P.  
 PR PR 20-OCT-2000; 2000US-0241786P.  
 PR PR 20-OCT-2000; 2000US-0241787P.  
 PR PR 20-OCT-2000; 2000US-0241808P.  
 PR PR 20-OCT-2000; 2000US-0241899P.  
 PR PR 20-OCT-2000; 2000US-0244782P.  
 PR PR 01-NOV-2000; 2000US-02446174P.  
 PR PR 08-NOV-2000; 2000US-02447474P.  
 PR PR 08-NOV-2000; 2000US-02447475P.  
 PR PR 08-NOV-2000; 2000US-02447476P.  
 PR PR 08-NOV-2000; 2000US-0244777P.  
 PR PR 08-NOV-2000; 2000US-0244782P.  
 PR PR 08-NOV-2000; 2000US-02448524P.  
 PR PR 08-NOV-2000; 2000US-0244525P.  
 PR PR 08-NOV-2000; 2000US-0244526P.  
 PR PR 08-NOV-2000; 2000US-0244527P.  
 PR PR 08-NOV-2000; 2000US-0244528P.  
 PR PR 08-NOV-2000; 2000US-0244610P.  
 PR PR 08-NOV-2000; 2000US-0244611P.  
 PR PR 08-NOV-2000; 2000US-0244613P.  
 PR PR 17-NOV-2000; 2000US-0249207P.  
 PR PR 17-NOV-2000; 2000US-0249213P.  
 PR PR 17-NOV-2000; 2000US-0249208P.  
 PR PR 17-NOV-2000; 2000US-0249211P.  
 PR PR 17-NOV-2000; 2000US-0249212P.  
 PR PR 17-NOV-2000; 2000US-0249218P.  
 PR PR 17-NOV-2000; 2000US-0249244P.  
 PR PR 17-NOV-2000; 2000US-0249245P.  
 PR PR 17-NOV-2000; 2000US-0249264P.  
 PR PR 17-NOV-2000; 2000US-0249365P.  
 PR PR 17-NOV-2000; 2000US-0249297P.  
 PR PR 17-NOV-2000; 2000US-0249299P.  
 PR PR 01-DEC-2000; 2000US-0250160P.  
 PR PR 05-DEC-2000; 2000US-0251030P.  
 PR PR 05-DEC-2000; 2000US-0251988P.

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.

XX Homo sapiens.

XX US2002132753-A1.

19-SEP-2002.

PP 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214885P.

PR 07-JUL-2000; 2000US-0216641P.

PR 07-JUL-2000; 2000US-0216830P.

PR 11-JUL-2000; 2000US-0217481P.

PR 11-JUL-2000; 2000US-0217498P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220967P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-022866P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229341P.

PR 01-SEP-2000; 2000US-0229344P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 08-SEP-2000; 2000US-0229518P.

PR 08-SEP-2000; 2000US-0231411P.

PR 21-SEP-2000; 2000US-0234222P.

PR 21-SEP-2000; 2000US-0234271P.

PR 23-SEP-2000; 2000US-0234997P.

PR 23-SEP-2000; 2000US-0235834P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0237040P.

PR 02-OCT-2000; 2000US-0237040P.

PR 02-OCT-2000; 2000US-0237042P.

PR 02-OCT-2000; 2000US-0237037P.

PR 20-OCT-2000; 2000US-0241785P.

PR 01-NOV-2000; 2000US-0241809P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237041P.

PR 17-NOV-2000; 2000US-0241299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251668P.

XX 08-DEC-2000; 2000US-0251869P.

PA (ROSE /) ROSEN C A.

PA (ROBE /) RUBEN S M.

PA (BARA /) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;

XX Kaser MR, Baughn MR, Au-Young J;

DR N-PSDB; ABX73534.

XX

PT New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or renal disorders.

XX

PT

PT

PT

PT

PS Claim 11; SEQ ID NO 1159; 402PP; English.

XX

CC The invention relates to human novel polypeptides and their associated polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma, blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABU5414-ABU5569 and ABU5748 represent human novel polypeptides of the invention

XX

XX

CC

XX



CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
 CC asthma. The products of the invention can be used for gene therapy or in  
 CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by  
 CC ABX97008-ABX97185

XX Sequence 1390 AA;

Query Match 82.6%; Score 500; DB 5; Length 1390;  
 Best Local Similarity 80.2%; Pred. No. 1.9e-5;  
 Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KPGRVNQOLQYHVKVMAKALWQHOPAWPPQPVDAVKGGLPDYHKKIKOPMDMGTTKRL 60  
 Db 55 KPRCTNQOLQYLRLVVKLTKLWQHOPAWPPQPVDAVKGGLPDYHKKIKOPMDMGTTKRL 114

Qy 61 ENNYWAASECMDNTMFTNCYLYNKPTDDIVLMAQTEKIFQVKASMP 111  
 Db 115 ENNYWAASECMDNTMFTNCYLYNKPTDDIVLMAQTEKIFQVKASMP 165

RESULT 10  
 ABB58985 standard; protein; 1937 AA.  
 XX  
 AC ABB58985;  
 AC  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster polypeptide SEQ ID NO 3747.  
 DE XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PBBE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 WPI; 2001-656860/75.  
 DR N-PSDB; ABIL03088.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX Disclosure: SEQ ID NO 3747; 21pp + Sequence Listing; English.  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABIL6176-ABIL0511), expressed DNA  
 CC sequences (ABIL01840-ABIL16175) and the encoded proteins (ABB5737-  
 CC ABB7202). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)  
 XX Sequence 1937 AA;

Query Match 82.3%; Score 498; DB 4;  
 Best Local Similarity 77.5%; Pred. No. 5.e-54;  
 Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ENNYWAASECMDNTMFTNCYLYNKPTDDIVLMAQTEKIFQVKASMP 111  
 Db 84 ENKYVZAKASECTEDFTMFSNCYLYNKPGDIVLMAQLEKLFNQKLSOMP 134

RESULT 12

AAV07114	AAV07114 standard; protein; 947 AA.	XX	ID ABR41326 standard; protein; 573 AA.
ID	AAV07114;	XX	XX
AC	AAV07114;	AC	ABR41326;
XX	02-JUL-1999 (first entry)	XX	DT 02-JUN-2003 (first entry)
DT		XX	Human DITHP intracellular signalling protein.
XX		XX	
DE	W09904265 Seq ID No: 685.	XX	Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis; cancer; cell proliferative disorder; autoimmune disorder; inflammatory disorder; infection; hormonal disorder; metabolic disorder; neurological disorder; gastrointestinal disorder; transport disorder; connective tissue disorder; drug screening; proteome analysis; gene therapy; antisense therapy; genotyping; transgenic animal; knock in; disease model; toxicological testing; transcript imaging; intracellular signalling.
XX	Cancer associated antigen; diagnosis; research; treatment; human; breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer; prostate cancer.	XX	
XX	Homo sapiens.	XX	XX
OS		OS	Homo sapiens.
XX	WO9904265-A2.	XX	XX
PN		PN	WO200297031-A2.
PD	28-JAN-1999.	PD	
XX	98W0-US014679.	XX	
PF	15-JUL-1998;	PF	
XX	98W0-US014679.	XX	
PR	17-JUL-1997;	PR	05-DEC-2002.
XX	97US-00886164.	XX	27-MAR-2002; 2002W0-US010056.
PR	10-OCT-1997;	PR	XX
XX	97US-0061199P.	PR	28-MAR-2001; 2001US-0278619P.
PR	10-OCT-1997;	PR	PR 29-MAR-2001; 2001US-0280067P.
XX	97US-0061765P.	PR	PR 29-MAR-2001; 2001US-0280068P.
PR	10-OCT-1997;	PR	PR 16-MAY-2001; 2001US-0291120P.
XX	97US-00948705.	PR	PR 17-MAY-2001; 2001US-0291829P.
PR	11-OCT-1997;	PR	PR 17-MAY-2001; 2001US-0291849P.
XX	97GB-00021697.	PR	PR 19-JUN-2001; 2001US-0299428P.
PR	22-JUN-1998;	PR	PR 20-JUN-2001; 2001US-0299776P.
XX	98US-00102322.	PR	PR 20-JUN-2001; 2001US-0300001P.
PA	(LUDWIG) LUDWIG INST CANCER RES.	XX	XX
XX	Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;	PA	(INCYT GENOMICS INC.
PI	PI	PI	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI	O'hare M, Obata Y, pfreundschuh M, Tureci O, Sahin U;	PI	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ansley SR,
XX	XX	PI	Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kiefield Y, Gersin BH;
WPI	WPI: 1999-132448/11.	XX	PI
XX	New isolated cancer associated nucleic acids and polypeptides - isolated using sera from cancer patients, used to develop products for the diagnosis, monitoring or treatment of cancers.	XX	PI
PT		XX	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PT		XX	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ansley SR,
XX		XX	Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kiefield Y, Gersin BH;
PS	Disclosure; Page 728-730; 787pp; English.	XX	PI
XX	The invention relates to a method for diagnosing a disorder characterized by expression of a human cancer associated antigen coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a biological sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the interaction between the agent and the NAM or the expression product as a determination of the disorder. The products and methods can be used in the diagnosis, monitoring, research, or treatment of conditions characterised by the expression of various cancer associated antigens.	XX	PI
CC	The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and lung cancer.	XX	PT
CC	Sequence 947 AA;	XX	PT
CC	Query Match 75.9%; Score 459; DB 2; Length 947;	XX	PT
CC	Best Local Similarity 73.0%; Pred. No. 2.1e-49;	XX	PT
CC	Matches 81; Conservative 16; Indels 0; Gaps 0;	XX	PT
Qy	1 KPGRTNQLQYLHKYKQAWKQAWPFPQPVDAVKGILPDPYHKIKOPMDGTIGRL 60	CC	CC
Db	24 KNGRLTQNQLQYLHKYKQAWKQAWPFPQPVDAVKGILPDPYHKIKOPMDGTIGRL 83	CC	CC
Qy	61 ENYYWAASSCQMDPNTMFTNCYLYNKPTDIVIMACTEKFQKVASMP 111	CC	CC
Db	84 ENKYAKASECIEDFTNTMSNCYLYNKPEQDIVMAQALEKLFMQLSQMP 134	CC	CC
CC	RESULTS 13	CC	CC
CC	ABR41326	CC	CC

CC mentioned above, as a source of antisense sequences, as a source of probes and primers, in genotyping and identification of individuals, in the generation of transgenic animal models of human disease or knock in humanized animals, in toxicological testing, and in transcript imaging. The present sequence represents a DITHP protein which has intracellular signalling activity. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

Sequence 573 AA;

Query Match 62.6%; Score 379; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2e-39; Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 DYHKIIKOPMDGTTIKRLENNYYWAASECMQDENTMFTNCIYNNKPTDDIVLMAQTLEX 101  
ID 5 DYHKIIKOPMDGTTIKRLENNYYWAASECMQDENTMFTNCIYNNKPTDDIVLMAQTLEX 64  
Db

Qy 102 IFLQKVASMP 111  
Db 65 IFLQKVASMP 74

RESULT 14  
AB68341; ID AB68341. standard; protein; 513 AA.  
XX ABB68341;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 31815.  
XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PR 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PCTE ) PE CORP NY.  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI: 2001-656860/75.  
DR N-PSDB; ABL12444.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA sequences (ABU1810-ABU16175) and the encoded proteins (ABBS5737-ABB207). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)  
Sequence 513 AA;

Query Match 46.5%; Score 281.5; DB 4; Length 513;  
Best Local Similarity 44.5%; Pred. No. 5.8e-27; Matches 49; Conservative 25; Mismatches 35; Indels 1; Gaps 1;

Qy 1 KPGRTVNTOLYHLKRVVMAIKWKGOPMDGTTIKRLENNYYWAASECMQDENTMFTNCIYNNKPTDDIVLMAQTLEX 60  
ID 36 RPGRTRNILEL-KSVNLNLWRPSYHFRHPYDSSVLSGVPDFYHANVKHPMDLSTIRKRL 94  
Db

Qy 61 ENNYWAASECMDNTMFTNCIYNNKPTDDIVLMAQTLEXKIFLQKAVSM 11.0  
ID 95 HNRYIWQWASEALDFKLIFDNCLLYNLESPVIAQAGKLLMEAFYMRMEI 144  
Db

RESULT 15  
ABR41589  
ID ABR41589 standard; protein; 330 AA.  
XX ABR41589;  
XX DT 02-JUN-2003 (first entry)  
XX DE Human DITHP antigen recognition protein.  
XX KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis; cancer; cell proliferative disorder; autoimmune disorder;  
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder; neurological disorder; gastrointestinal disorder; transport disorder; connective tissue disorder; drug screening; proteome analysis;  
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in; disease model; toxicological testing; transcript imaging; antigen recognition.  
XX DE Homo sapiens.  
XX OS Homo sapiens.  
XX PN WO200297031-A2.  
XX PD 05-DEC-2002.  
XX PF 27-MAR-2002; 2002WO-US010056.  
XX PR 26-MAR-2001; 2001US-0279619P.  
PR 29-MAR-2001; 2001US-0280667P.  
PR 29-MAR-2001; 2001US-0280668P.  
PR 16-MAY-2001; 2001US-0291820P.  
PR 17-MAY-2001; 2001US-0291829P.  
PR 17-MAY-2001; 2001US-0291849P.  
PR 19-JUN-2001; 2001US-0299428P.  
PR 20-JUN-2001; 2001US-0299476P.  
PR 20-JUN-2001; 2001US-0300001P.  
XX (INCYT-) INCYTE GENOMICS INC.  
XX PA DR WPI: 2003-1229518/12.  
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Daffo GE, Hillman JL, Yu JY, Tran O, Yap PE, Amshay SR;  
PI Daugertey SC, Dam TC, Liu TF, Nguyen DA, Kliefeld Y, Gersin EH;  
PI Perata CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
PI Flores V, Maraha R, Lo A, Lan YY, Urashka ME;  
XX DR N-PSBB; ACC46537.  
XX The invention relates to novel human diagnostic and therapeutic polynucleotides which specifically binds to a polypeptide encoded by human diagnostic and therapeutic polynucleotide, and to induce antibodies.  
CC Claim 27; SEQ ID NO 1124; 591PP; English.  
CC The invention relates to novel human diagnostic and therapeutic polynucleotide useful for identifying PT test compound which specifically binds to a polypeptide encoded by human PT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
CC XX  
CC Novel human diagnostic and therapeutic polypeptide useful for identifying PT test compound which specifically binds to a polypeptide encoded by human PT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
CC XX  
CC The invention also relates to the dithp cDNA CC sequences of the invention; recombinant vectors, host cells and CC sequences of the invention; recombinant vectors, host cells and CC

transgenic organisms comprising a dithp nucleic acid sequence; the recombinant production of DITHP proteins; antibodies specific for DITHP proteins; microarrays comprising dithp nucleic acid sequences; methods of detecting dithp nucleic acid and protein sequences; methods of screening for compounds which specifically bind a DITHP protein; and methods of assessing the toxicity of test compounds using a dithp probe. Dithp nucleic acid sequences and DITHP proteins may be used in the diagnosis of a wide variety of conditions including cancer and other cell proliferative disorders; autoimmune or inflammatory disorders; bacterial, viral, fungal or parasitic infections; hormonal disorders; metabolic disorders; neurological disorders; gastrointestinal disorders; transport disorders; and connective tissue disorders. They may also be used to screen for modulators of protein activity or gene expression. DITHP proteins can additionally be used in analysis of the proteome of a tissue or cell type and to induce antibodies. The dithp nucleic acids are additionally useful in somatic or Germline gene therapy of the disorders mentioned above, as a source of antisense sequences, as a source of probes and primers in genotyping and identification of individuals, in the generation of transgenic animal models of human disease or knock in humanised animals, in toxicological testing, and in transcript imaging. The present sequence represents a DITHP protein which has antigen recognition activity. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

SQ Sequence 330 AA;

	Query Match	Score	DB	Length
Qy	1 KPGRVINOLQYLHKVVMKALWKRHQ - FAVFPRQPYDAYKLGLPDYHKKIKQPMMDGTTKR	44.5%	269	6
Db	153 KKGKLUSEQQLRHCGNLKLLSKVKAAYAWPFYKEVDASALGLHDYHDIKHMDSLTVKR	46.0%	25	58
Qy	59 RLNNYYWAASECMODFNTMFTNCYVYNKTFDDTFLVMAQTLKFLQKVASMP	38	Indels	212
Db	213 KMENRDYRDQEFQAAVRLMFSNCYKYNPPDHDVAMARKLQDVPEFRYAVMP	38	Mismatches	111
		1	Gaps	265

Search completed: July 9, 2004, 13:06:14  
 Job time : 58 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

Om protein - protein search, using sw model

Run on: July 9, 2004, 13:04:05 ; Search time 21 Seconds

(without alignments)  
 508,441 Million cell updates/sec

Title: US-09-784-553C-19  
 Perfect score: 605  
 Sequence: 1 KGRVTNQIQLHKKVAKAL.....IVLMAQTLKIFLQYKVASMP 111

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : PIR 78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	605	100.0	754	2	A56619	female sterile hom RING3 kinase - chicken
2	594	98.2	733	2	T28145	female sterile hom RING3 kinase - chicken
3	498	82.3	2038	2	A43742	hypothetical protein
4	296	48.9	1250	2	T22945	hypothetical protein
5	279	46.1	1087	2	T22847	kinase-like protein
6	221	36.6	703	2	T48560	transcription factor
7	215	35.5	578	2	T40384	probable RING3 protein
8	204	33.7	400	2	T00472	hypothetical protein
9	201	33.2	461	2	D96757	PCAF protein - human
10	198	32.7	832	1	S71758	bromodomain protein
11	195	32.3	374	2	T33328	bromodomain protein
12	195	32.2	678	2	T49984	bromodomain protein
13	190	31.5	361	2	T42517	hypothetical protein
14	188	31.1	766	2	A86198	F2H15.2 protein -
15	187	30.9	440	2	H86312	CREB-binding protein
16	184	30.5	2441	2	S39161	hypothetical protein
17	183	30.3	638	2	S67605	transcription factor
18	182	30.1	439	1	S28051	transcription factor
19	181	30.0	2414	2	A54277	transcription factor
20	180	29.8	2440	2	S39162	transcription factor
21	179	29.7	586	2	T47620	histone acetyltransferase
22	176	29.1	405	2	T21433	hypothetical protein
23	176	29.1	452	2	T21335	hypothetical protein
24	176	29.1	510	2	T21410	hypothetical protein
25	176	29.1	686	2	S5955	bromodomain protein
26	172	28.4	454	2	T37933	transcription activator
27	168	27.9	769	2	E96613	hypothetical protein
28	166	27.4	3190	2	T13828	CREB-binding protein
29	162	26.8	2027	2	S60123	hypothetical protein

#### RESULT 1

A56619  
 female sterile homeotic (fsh) homolog RING3 - human

C:Species: Homo sapiens (man)

C:Accession: A56619; S18860; S40781

R:Beck, S.; Hanson, I.; Kelly, A.; Pappin, D.J.; Trowsdale, J.

DNA Seq. 2, 203-210, 1992

A:Title: A homologue of the *Drosophila* female sterile homeotic (fsh) gene in the class I

A:Reference number: A56619; PMID:92329974; PMID:1352711

A:Cross-references: EMBL:X62083; NID:931471; PID:931472; EMBL:MB0513; N

A:Note: sequence extracted from NCBI backbone (NCBIP:108781);

A:Accession: A56619

A:Status: preliminary

A:Gene: RING3

A:Keywords: duplication

C:Genetics: unassigned bromodomain proteins; bromodomain homology

F:52-109/Domain: bromodomain homology <BR>

F:325-382/Domain: bromodomain homology <BR>2>

Query Match

100.0%; Score: 605;

DB 2;

Length: 754;

Best Local Similarity

100.0%; Pred. No. 7.6e-58;

Mismatches

0;

Indels

0;

Gaps

0;

Qy

1 KPGRTINQYQIHKYQMKALWQFAWPRPQPVDAVKGDPYHKKIKOPMDMGTTKRL 60

24 KPGRTINQYQIHKYQMKALWQFAWPRPQPVDAVKGDPYHKKIKOPMDMGTTKRL 83

Dy

61 ENNYWAASEMQDNTMFTNCFYKNCPKPDDIVLMAQLEKIFLQKVASMP 111

Qy

64 ENNYWAASEMQDNTMFTNCFYKNCPKPDDIVLMAQLEKIFLQKVASMP 134

Dy

RESULT 2

T28145  
 RING3 kinase - chicken

C:Species: Gallus gallus (chicken)

C:Accession: T28145

R:Maile, S.; Kaufman, J.; Beck, S.

DNA sequence, May 1998

submitted to the EMBL Data Library

A:Description: DNA sequencing and analysis of the chicken major histocompatibility complex

A:Reference number: 220475

A:Accession: T28145

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-733 <NL>

A:Cross-references: EMBL:AL023516; NID:e1292539; PID:e1292549; A:Experimental source: clone CB12

A:Accession: clone CB12

#### ALIGNMENTS

C;Genetics:  
 A;Gene: RING3  
 A;Map position: 16  
 A;Introns: 64/3; 110/3; 158/1; 227/3; 351/3; 394/3; 479/3; 546/2; 650/1; 691/1  
 C;Superfamily: unassigned bromodomain proteins; bromodomain homology  
 P;52-109/Domain: bromodomain homology <BRO1>  
 P;323-380/Domain: bromodomain homology <BRO2>

Query Match 98.2%; Score 594; DB 2; Length 733;  
 Best Local Similarity 97.3%; Pred. No. 1.2e-56;  
 Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 A;Genetics: C;Genetics:  
 A;Map position: X  
 A;Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3  
 C;Superfamily: bromodomain homology  
 P;307-364/Domain: bromodomain homology <BRO1>  
 P;619-676/Domain: bromodomain homology <BRO2>

Query Match 61 ENNYWAASSCMQDNTMFNCYIYNKPTDDIVLMAQTLKFLQKVASMP 111  
 Best Local Similarity 46.8%; Pred. No. 7e-24;  
 Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;  
 C;Accession: A43742; B41742  
 C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 20-Sep-1999  
 R;Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; David, I.B.  
 Dev. Biol. 134, 246-257, 1989  
 A;Title: The Drosophila fish locus, a maternal effect homeotic gene, encodes apparent membrane sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)  
 N;Alternate names: membrane protein fish, 205K  
 N;Contains: female sterile homeotic protein, 110K  
 C;Species: Drosophila melanogaster  
 C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 20-Sep-1999  
 A;Reference: A43742; B41742  
 A;Accession: A43742; MUID:89276730; PMID:2567251  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 12038 <HAY>  
 A;Cross-references: EMBI:MM23221; NID:9157452; PID:9157453  
 A;Accession: B43742  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 11106 <HA2>  
 A;Cross-references: EMBI:MM23222  
 A;Gene: fsh  
 A;Cross-references: FlyBase:FBgn0004656  
 C;Superfamily: unassigned bromodomain proteins; bromodomain homology  
 C;Keywords: alternative splicing; membrane protein  
 P;11106/Domain: female sterile homeotic protein, 205K #status predicted <MA2>  
 P;5-116/Domain: bromodomain homology <BRO1>  
 P;503-560/Domain: bromodomain homology <BRO2>

Query Match 82.3%; Score 498; DB 2; Length 2038;  
 Best Local Similarity 77.5%; Pred. No. 1.2e-45;  
 Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;  
 C;Accession: T22845  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1993 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 C;Accession: T22845  
 R;White, S.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Meviss, H.W.; Rudd, S.; Len  
 R;Bevan, M.

submitted to the EMBL Data Library, February 1996  
 A;Reference number: Z19625  
 A;Accession: T22845  
 A;Status: preliminary; translated from GB/EMBL/DBBJ  
 A;Molecule type: DNA  
 A;Residues: 1-125 <WIL>  
 A;Cross-references: EMBI:Z69616; PIDN:CAA93473.1; GSPDB:GN00028; CESPDB:Z57C7.1a  
 A;Experimental source: clone F57C7  
 C;Genetics:  
 A;Map position: X  
 A;Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3  
 C;Superfamily: bromodomain homology  
 P;307-364/Domain: bromodomain homology <BRO1>  
 P;619-676/Domain: bromodomain homology <BRO2>

Query Match 48.9%; Score 296; DB 2; Length 1250;  
 Best Local Similarity 46.8%; Pred. No. 7e-24;  
 Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;  
 C;Accession: T22847  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1993 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 R;White, S.  
 Submitted to the EMBL Data Library, February 1996  
 A;Reference number: Z19625  
 A;Accession: T22847  
 A;Status: preliminary; translated from GB/EMBL/DBBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1087 <WIL>  
 A;Cross-references: EMBI:Z69646; PIDN:CAA93475.1; GSPDB:GN00028; CESPDB:Z57C7.1b  
 A;Experimental source: clone F57C7  
 C;Genetics:  
 A;Map position: X  
 A;Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1  
 C;Superfamily: bromodomain homology  
 P;307-364/Domain: bromodomain homology <BRO1>  
 P;579-636/Domain: bromodomain homology <BRO2>

Query Match 46.1%; Score 279; DB 2; Length 1087;  
 Best Local Similarity 44.1%; Pred. No. 4.3e-22;  
 Matches 49; Conservative 24; Mismatches 38; Indels 0; Gaps 0;  
 C;Accession: T22845  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1993 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 C;Accession: T22845  
 R;White, S.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Meviss, H.W.; Rudd, S.; Len  
 R;Bevan, M.

RESULT 6  
 T48600  
 kinase-like protein - *Arabidopsis thaliana*  
 N;Alternate names: protein F18022.60  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000  
 C;Accession: T48600  
 R;Bevan, M.

Query Match 61 ENNYWAASSCMQDNTMFNCYIYNKPTDDIVLMAQTLKFLQKVASMP 111  
 Best Local Similarity 46.8%; Pred. No. 7e-24;  
 Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;  
 C;Accession: T22847  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1993 #sequence\_revision 15-Oct-1999 #text\_change 02-Sep-2000  
 R;White, S.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Meviss, H.W.; Rudd, S.; Len  
 R;Bevan, M.

RESULT 6  
 T48600  
 kinase-like protein - *Arabidopsis thaliana*  
 N;Alternate names: protein F18022.60  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000  
 C;Accession: T48600  
 R;Bevan, M.

Submitted to the Protein Sequence Database, April 2000

A; Reference number: 224493

A; Accession: T48600

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-703 &lt;BEV&gt;

A; Cross-references: EMBL:DB

A; Experimental source: cultivar Columbia; BAC clone F18O22

C; Genetics:

A; Map position: 5

A; Introns: 370/3; 387/1; 423/1; 461/3; 484/1; 511/1; 549/3; 572/1; 619/3

A; Note: F18O22\_60

C; Superfamily: bromodomain homology &lt;BRO&gt;

F;157-214/Domain: bromodomain homology &lt;BRO&gt;

A; Reference number: 224160

A; Accession: T00472

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-400 &lt;ROU&gt;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;STO&gt;

A; Cross-references: GB:AE002093; PID:G3033386; PIDN: AAC12830\_1; GSPDB:GN00139

C; Genetics:

A; Gene: At2g34900; F19I3.13

A; Map position: 2

A; Introns: 6/3; 49/3; 120/3; 146/3; 218/3; 251/3; 271/2; 335/3

C; Superfamily: bromodomain homology

F;130-187/Domain: bromodomain homology &lt;BRO&gt;

A; Reference number: 221962

A; Accession: T40984

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-578 &lt;LYN1&gt;

A; Cross-references: EMBL:AL049559; PIDN: CAB40169\_1; PID:G4581512; GSPDB:GN

A; Experimental source: strain 972h-; cosmid c1450

A; Authors: Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

A; Position: submitted to the EMBL Data Library, March 1999

A; Reference number: 221904

A; Accession: T41225

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-111 &lt;LYN2&gt;

A; Cross-references: EMBL:AL049644; PID:G4578667; PIDN: CAB41059\_1; PID:G4678679; GSPDB:GN

A; Genetics: &lt;C191&gt;

A; Gene: SPDB:SPCC1191.13

A; Map position: 1

C; Superfamily: bromodomain homology &lt;BRO1&gt;

F;109-166/Domain: bromodomain homology &lt;BRO2&gt;

F;219-336/Domain: bromodomain homology &lt;BRO2&gt;

A; Reference number: 221904

A; Accession: T41225

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-215 &lt;ROU&gt;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C; Accession: D96757

C; Authors: Rizzoli, M.; Sakano, H.; Rowley, T.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

C; Position: 1

A; Description: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A; Reference number: A86111; PMID:11016719

A; Accession: D96757

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-461 &lt;STO&gt;

A; Cross-references: GB:AE005173

A; Cross-references: GB:AE00141

Db 307 QSLXNKNEYSTLDEFEISDILLMENNCFTYNNPDPGTPVHMGROLENVKEKNEARP 361

RESULT 8

T00472 probable RING3 protein [imported] - Arabidopsis thaliana

N; Alternative names: hypothetical protein F19I3.13

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-400 &lt;ROU&gt;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;STO&gt;

A; Cross-references: GB:AE002093; PID:G3033386; PIDN: AAC12830\_1; GSPDB:GN00139

C; Genetics:

A; Gene: At2g34900; F19I3.13

A; Map position: 2

A; Introns: 6/3; 49/3; 120/3; 146/3; 218/3; 251/3; 271/2; 335/3

C; Superfamily: bromodomain homology

F;130-187/Domain: bromodomain homology &lt;BRO&gt;

Query Match 36.6% Score 221.5; DB 2; Length 703;

Best Local Similarity 41.5%; Pred. No. 4.7e-16; Indels 3; Gaps 1;

Matches 44; Conservative 15; Mismatches 44; Gaps 0;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;ROU&gt;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;STO&gt;

A; Cross-references: GB:AE005173

A; Cross-references: GB:AE00141

Db 307 QSLXNKNEYSTLDEFEISDILLMENNCFTYNNPDPGTPVHMGROLENVKEKNEARP 361

RESULT 9

T40984 probable RING3 protein [imported] - Arabidopsis thaliana

N; Alternative names: protein SPCC191.13

C; Species: Schizosaccharomyces pombe

C; Date: 03-Dec-1999 #sequence\_revision 03-Dec-2000

C; Accession: T40984; T41225

C; Authors: Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

C; Position: submitted to the EMBL Data Library, March 1999

A; Reference number: 221962

A; Accession: T40984

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-578 &lt;LYN1&gt;

A; Cross-references: EMBL:AL049559; PIDN: CAB40169\_1; PID:G4581512; GSPDB:GN

A; Experimental source: strain 972h-; cosmid c1450

A; Authors: Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

A; Position: submitted to the EMBL Data Library, March 1998

A; Reference number: 221904

A; Accession: T41225

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-111 &lt;LYN2&gt;

A; Cross-references: EMBL:AL049644; PID:G4578667; PIDN: CAB41059\_1; PID:G4678679; GSPDB:GN

A; Genetics: &lt;C191&gt;

A; Gene: SPDB:SPCC191.13

A; Map position: 1

C; Superfamily: bromodomain homology &lt;BRO1&gt;

F;109-166/Domain: bromodomain homology &lt;BRO2&gt;

F;219-336/Domain: bromodomain homology &lt;BRO2&gt;

Query Match 35.5% Score 215; DB 2; Length 578;

Best Local Similarity 40.9%; Pred. No. 1.9e-15; Indels 4; Gaps 2;

Matches 47; Conservative 15; Mismatches 49; Gaps 49;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;ROU&gt;

A; Cross-references: EMBL:AC004238; PID:G3033373; PID:G3033386

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C; Accession: T00472; CB4762

C; Authors: Rounseley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, J.

C; Position: submitted to the EMBL Data Library, April 1998

A; Description: Arabidopsis thaliana chromosome II BAC F19I3 genomic sequence.

A; Reference number: Z:4160

A; Accession: T00472

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-400 &lt;STO&gt;

A; Cross-references: GB:AE005173

A; Cross-references: GB:AE00141

Db 307 QSLXNKNEYSTLDEFEISDILLMENNCFTYNNPDPGTPVHMGROLENVKEKNEARP 361

RESULT 8

T40984 probable RING3 protein [imported] - Arabidopsis thaliana

N; Alternative names: protein SPCC191.13

C; Species: Schizosaccharomyces pombe

C; Date: 03-Dec-1999 #sequence\_revision 03-Dec-2000

C; Accession: T40984; T41225

C; Authors: Lyne, M.; Raj

C;Genetics:  
A;Gene: T18K17.19  
A;Map position: 1

Query Match 33.2%; Score 201; DB: 2; Length 451;  
Best Local Similarity 42.9%; Pred. No. 5e-14; Indels 0; Gaps 0;  
Matches 39; Conservative 14; Mismatches 38; Indels 0; Gaps 0;

Qy 16 VMKALWKHOFAWPFRQPVDAVKLGLPDPYHKIKQPMDMGTTIKRNLIKQPMDMGTTIKRNL 61  
Db 126 LLTKLWKHKSSWIFNTPVDAVKLGLPDPYHKIKQPMDMGTTIKRNL 185

Qy 76 NTMFNCYIYKPTDDIVMAQTLKEFLQK 106  
Db 186 RLTFNNAMLNPNVGHDVYHAEILLNLFEKK 216

RESULT 10  
T19984  
P/CAF protein - human  
N;Alternate names: CREBBP-associated factor  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 29-Sep-1999  
C;Accession: S1788  
C;Description: T19984  
R;Yang, X.J.; Oryzko, V.V.; Nishikawa, J.; Howard, B.H.; Nakatani, Y.  
Nature 382, 319-324, 1996  
A;Title: A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A.  
A;Reference number: S71788; MUID:96300317; PMID:8684459  
A;Status: Preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-832 <YAN>  
A;Cross-references: EMBL:U57317; NID:91491936; PIDN: AAC50890.1; PID:91491937  
C;Genetics:  
A;Gene: GDB:CAF  
A;Cross-references: GDB:9864231; OMIM:602303  
A;Map position: 3p24-3p4  
C;Superfamily: human P/CAF Protein; bromodomain homology  
E;7.28-8.03/Domain: bromodomain homology <BRO1>

Query Match 32.7%; Score 198; DB: 1; Length 832;  
Best Local Similarity 37.4%; Pred. No. 2.1e-13;  
Matches 40; Conservative 18; Mismatches 47; Indels 2; Gaps 1;

Qy 1 KPGRTVNOLOYLHKVVMKALWKHOFAWPFRQPVDAVKLGLPDPYHKIKQPMDMGTTIKRNL 60  
Db 720 KEPRDPDQLVSTLKSILQQKSHSSAWPMEPVKRTE--APGYEVRSRPMQLKTMNSRL 777

Qy 61 ENNYWAAASSCMQDFNTMFTNCYIYKPTDDIVMAQTLKEFLQK 107  
Db 778 KNRYVSKLFLMADLQRVFTNCKEYNAPESYYRKCANILEKEFFESKI 824

RESULT 11  
T33328  
hypothetical protein F13C5.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 02-Sep-2000  
C;Accession: T33328  
R;Kehldmann, B.; Hawkins, J.; Gilliam, B.  
.submitted to the EMBL Data Library, July 1998  
A;Description: The sequence of C. elegans cosmid F13C5.  
A;Accession: T33328  
A;Status: Preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-374 <WOH>  
A;Cross-references: EMBL: D89531; PIDN: AAC64610.1; GSPDB: GN00028; CESP: F13C5.2  
A;Experimental source: strain PR45  
C;Genetics:  
A;Gene: CESP:F13C5.2  
A;Map position: X  
A;Introns: 25/3; 135/3; 189/3; 313/3;  
A;Introns: 1374 <WOH>  
A;Cross-references: EMBL: D89157; PID: 91749521; PIDN: BAA13819.1; PID: 91749522  
A;Experimental source: strain PR45  
C;Superfamily: bromodomain homology  
F;64-121/Domain: bromodomain homology <BRO>

Query Match 31.5%; Score 190.5; DB: 2; Length 361;  
Best Local Similarity 40.4%; Pred. No. 5.2e-13;





Scoring table:	BLOSUM62	Gapop 10.0 , Gapext 0.5	ALIGNMENTS
Searched:	141681 seqs,	52070155 residues	RESULT 1
Total number of hits satisfying chosen parameters:	141681		BRD2_HUMAN STANDARD; PRT; 801 AA.
Minimum DB seq length: 0			ID BRD2_HUMAN; P25440; Q00699; Q00700; Q15310; Q96904;
Maximum DB seq length: 20000000000			AC P25440; Q00699; Q00700; Q15310; Q96904;
Post-processing: Minimum Match 0%			DT 01-MY-1992 (Rel. 22, Created).
Maximum Match 100%			DT 16-OCT-2001 (Rel. 40, Last sequence update).
Listing first 45 summaries			DT 28-FEB-2003 (Rel. 41, Last annotation update).
Database :	SwissProt_42:*		DE Bromodomain-containing protein 2 (RING3 protein) (O27.1.1).
Scoring table:	BLOSUM62	Gapop 10.0 , Gapext 0.5	GN BRD2 OR RING3 OR KIAA9001.
Perfect score:	605		OS Homo sapiens (Human).
Sequence:	1 KPGRTVNOLQYLHKVVMKAL.....IVIMATQLEKIFLQKVASMP 111		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo; OC NCBITaxID=9606;
			OX RN [1]
			SEQUENCE FROM N.A.
			RP TISSUE=T-cell;
			RC RX MEDLINE=9332994; PubMed=1352711;
			RT Beck S., Hanson I., Kelly A., Papin D.J.C., Trowsdale J.,
			RT "A homologue of the <i>Drosophila</i> female sterile homeotic (fsh) gene in the class II region of the human MHC.";
			RT DNA Seq. 2:203-210(1992).
			RN [2]
			SEQUENCE TO N-TERMINUS.
			RP MEDLINE=9337636; PubMed=8781126;
			RC Thorge K.L., Abdulla S., Kauiman J., Trowsdale J., Beck S.;
			RA Phylogeny and structure of the RING3 gene.";
			RT Immunogenetics 44:391-396(1996).
			RL RN [3]
			SEQUENCE FROM N.A.
			RP TISSUE=Bone marrow;
			RC Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y., RA Sato S., Nagase T., Seki T., Ishikawa K.-I., Tabata S.;
			RA Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
			RL RN [4]
			SEQUENCE FROM N.A.
			RP Thorpe K., Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
			RL CC !- SUBCELLULAR LOCATION: Nuclear (Potential).
			CC CC !- SIMILARITY: Contains 2 bromodomains.
			CC CC !- SIMILARITY: Contains 1 ET domain.
			CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation
			CC the European Bioinformatics Institute. There are no restrictions on its
			CC use by non-profit institutions as long as its content is in no way
			CC modified and this statement is not removed. Usage by and for commercial
			CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
			CC or send an email to license@isb-sib.ch).
			CC
			DR EMBL X62083; QAA43996.1;
			DR EMBL P06113; PAA68880.1; ALT_INIT.
			DR EMBL M82040; PAA07841.1;
			DR EMBL D42040; PAA07841.1;
			DR EMBL 284497; AAC69991.1;
			DR EMBL 296104; AAC69989.1;
			DR EMBL X96570; AAC69989.1;
			DR HSSP; Q92831; IB91.



use by non-profit institutions as long as its content is in no way modified and this license agreement (see <http://www.ist-sib.ch/announce/> or send an email to [license@ist-sib.ch](mailto:license@ist-sib.ch)).

RA	Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gellites S., Goble A., Hamlin N., Harris D., Hedges S., Hedges G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hodges G., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Rabinowitch E., Rutherford K., Buttner S., Saunders D., Seeger K., Simmonds M., Squares R., Stevens K., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Whitehead S., Weltjens I., Vanstreels B., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritsch C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Punelle B., Goffau A., Cadieu E., Dreano S., Gloux S., Leilaire V., Mottier S., Gilbert F., Ayres S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forisburg S.L., Cerrutti L., Lowe T., McCombe W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P., "The Genome sequence of Schizosaccharomyces pombe," Nature 415:871-880 (2002).	RA	PCAF HUMAN STANDARD: PRT: 832 AA. AC Q92831; DT 28-FEB-2003 (Rel. 41, Created) DT 28-FEB-2003 (Rel. 41, Last sequence update) DT 15-MAR-2004 (Rel. 43, Last annotation update) DT P30/EP-associated factor (EC 2.3.1.-) (P/CAF) (Histone acetylase PCAF). DE PCAF. GN Homo sapiens (Human). OC Eukaryota; Metzoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. NCBI_TAXID:9606; RN [1] SEQUENCE FROM N.A., FUNCTION, SUBUNIT, TISSUE SPECIFICITY, AND INTERACTION WITH P300 AND CBP. RP TISSUE=Liver; RC MEDLINE=96100317; PubMed=8684459; RA Yang X.-J.; Ogrzyko V.V.; Nishikawa J.-I.; Howard B.H.; Nakatani Y.; "A P300/CBP-associated factor that competes with the adenoviral oncoprotein P1A," Nature 382:319-324 (1996). RN [2] REVISIONS. RP TISSUE=Liver; RC MEDLINE=96100317; PubMed=8684459; RA Nakatani Y.; "A P300/CBP-associated factor that competes with the adenoviral oncoprotein P1A," Nature 382:319-324 (1996). RN [3] Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases. RN [4] ENZYMATIC ACTIVITY. RP MEDLINE=97100944; PubMed=8945521; RA Ogrzyko V.V.; Schiltz R.L.; Russanova V.; Howard B.H.; Nakatani Y.; "The transcriptional coactivators p300 and CBP are histone acetyltransferases," Cell 87:933-959 (1996). RN STRUCTURE BY NMR OF 715-832, AND MUTAGENESIS OF VAL-752; TYR-760; TYR-802 AND TYR-809. RP TISSUE=Liver; RC MEDLINE=99192086; PubMed=10365964; RA Dhalluin C., Carlson J.E., Zeng L., He C., Aggarwal A.K., Zhou M.-M.; "Structure and ligand of a histone acetyltransferase bromodomain," Nature 399:491-496 (1999). RN [-] FUNCTION: Functions as a histone acetyltransferase (HAT) to promote transcriptional activation. Has significant histone acetyltransferase activity with core histones (H3 and H4), and also with nucleosome core particles. Inhibits cell-cycle progression and counteracts the mitogenic activity of the adenoviral oncoprotein E1A. CC [-] SUBUNIT: Interacts with SIRT1 (By similarity). Interacts with CC p300 and CBP. CC [-] SUBCELLULAR LOCATION: Nuclear (By similarity). CC [-] TISSUE SPECIFICITY: Ubiquitously expressed but most abundant in heart and skeletal muscle. CC [-] SIMILARITY: Contains 1 bromodomain. CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). CC EMBL; AL394713; CAC05484.1; -. CC EMBL; D89157; BAA1819.1; -. CC PIR; T42517; T42517.2. CC GeneBank Spombe; SPAC531.02; -. CC InterPro; IPR001487; Bromodomain. CC Pfam; PF00439; bromodomain_2. CC PRINTS; PR00503; bromodomain. CC SMART; SM00297; BRON_2. CC PROSITE; PS00633; bromodomain_1; 1. CC PROSITE; PS50014; bromodomain_2; 2. CC Hypothetical protein; bromodomain; repeat. FT DOMAIN 246 318 bromodomain_1. FT DOMAIN 408 480 bromodomain_2. SQ SEQUENCE 727 AA; 81031 MW; 5CPA7384CB6FF05 CRC64; Query Match 2 PGRVITNQLOVYIHKVYKPAWPFQCPDVAKLGPDYHKIKPQMDNGTIKKRL 61 Best Local Similarity 33.5%; Score 202.5; DB 1; Length 727; Matches 37; Conservative 29; Mismatches 43; Indels 1; Gaps 1; YQ 2 PGRVITNQLOVYIHKVYKPAWPFQCPDVAKLGPDYHKIKPQMDNGTIKKRL 61 Db 228 PPMTEQHKYTH-AMLRQFRGRDSSPFAPVDPVQKNIQDPTIKNPDLGTMQRKFS 286 YQ 62 NNYWAAESECMQDFNTMFTMFTYIYKPTDDIVMAQTLKIFQKVASMP 111 Db 287 SGVYSSAQHFFDDMNLMSNCFLYNGTTESSVGVNGKSNLQATPFERQLQKLP 336 RESULT 6 PCAF_HUMAN
----	---	----	---



Db	1099	LEALYRQDPESLPFRQPVDPOLLGIPDYFIDIVKQMDLSPIKRLDQYQEPHQYVYDDV	1158	"Sequence analysis of a 10.5 kb DNA fragment from the yeast chromosome VII reveals the presence of three new open reading frames and of a <i>TANAThr</i> gene.";
Qy	76	NTMFNCYIYNKPKPTDIVLMAQTLKIFQKV 107		Yeast 13: 369-372 (1997).;
Db	1159	RLMFNNAWLYNPKTSRKYKFCSKLAEVPEEAI	1190	[5]
				ASSOCIATION WITH ADA2.
RESULT 8				RP MEDLINE=95045371; PubMed=7957049;
GCN5 YEAST				RT ADA2: putative transcriptional adaptors.;
Q03333; ID STANDARD;				RT RT EMBO J. 13:4807-4815(1994).
AC PRT; 439 AA.				RT RN RT "Functional similarity and physical association between GCN5 and ADA2; ADA3 AND TRAI.
DT 01-OCT-1993 (Rel. 27, Created)				RT RP MEDLINE=99102959; PubMed=9885573;
DT 01-OCT-1993 (Rel. 27, Last sequence update)				RT RA Grant P.A., Schiltz D., Pray-Grant M.C., Yates J.R. III, Workman J.L.;
DT 15-MAR-2004 (Rel. 43, Last annotation update)				RT RA "The ATM-related cofactor Tral is a component of the purified SAGA complex.";
DE Histonetransferase GCN5 (EC 2.3.1.48).				RT RL Mol. Cell 2:863-867(1998).
GN GCN5 OR ADA4 OR YGR252W.				RT RN [7]
OS Saccharomyces cerevisiae (Baker's yeast).				RP REVIEW
OC Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				RX MEDLINE=20400931; PubMed=10940244;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces;				RA Dyda F., Klein D.C., Hickman A.B.;
OX NCBI_TAXID:4932;				RT RT "GCN5-related N-acetyltransferases: a structural overview.";
RN				RL Annu. Rev. Biophys. Biomol. Struct. 29:81-103(2000).
RP SEQUENCE FROM N.A.				RA RN [8]
RX MEDLINE=9301009; PubMed=1399595;				RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 99-262.
RA Georgiaopoulos T., Thiresos G.;				RX MEDLINE=99362688; PubMed=10430873;
RT "Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription.";				RA Trievell R.C., Rojas J.R., Stern D.E., Venkataraman R.N., Wang L., Zhou J., Allis C.D., Berger S.L., Marinowitz R.J.
RL EMBO J. 11:4145-4152(1992).				RT RT "Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator.";
RN				RL Proc. Natl. Acad. Sci. U.S.A. 96:931-936(1999).
RP SEQUENCE FROM N.A.				CC -!- FUNCTION: Acetylates Lys-14 of histone H3. Also acetylates Lys-8 and Lys-16 of histone H4 with a lower preference. Acetylation of histones gives a specific tag for epigenetic transcription activation. Operates in concert with certain DNA-binding transcriptional activators such as GCN4 or HAP2/3/4. Acts via the formation of large multiprotein complexes that modify the chromatin.
RX MEDLINE=9313265; PubMed=9168696;				CC CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
RA Tettelin H., Agostoni Carboni M.L., Albermann K., Albers K., Arroyo K., Backes U., Barreiros T., Bertani I., Björnsson A.-J., Brueckner M., Bruschi C.V., Cariogiani G., Castagnoli L., Cerdan E., Clemente M.L., Coblenz A., Cogliavina M., Coissac E., Del Bino S., Delius H., Della P., Dierler D., Dierig P., Dujon B., Duraid P., Enright K.-D., Eraso P., Escrivano V., Faiola L., Farthmann B., Feroli P., Feuermann M., Frontali L., Garcia-Gonzalez M., Hebling U., Hernandez K., Heumann K., Hofmann B., Indge K.J., James C.M., Klima R., Hilger F., Hofmann B., Krammer W., Launay G., Leutheuer H., Louis B.J., Maillet B., Marconi A., Martegani E., Mazon M.J., Mazzoni C., McReynolds A.D.K., Melchiorreto P., Newes H.-W., Mienkova O., Mueller-Auer S., Nawroki A., Netter P., Neu R., Nombela C., Oliver S.G., Panzeri L., Paoluzi S., Plevani P., Porteille D., Portillo F., Potier S., Purnelle B., Rieger M., Riles L., Rinaldi T., Robben J., Rodriguez Pousada C., Rodriguez-Belmonte E., Rodriguez-Torres A.M., Rose M., Ruzzi M., Salio M., Sanchez-Perez M., Schaefer B., Schaefer M., Schärfe M., Schmidt-Haini T., Schreer A., Skalak J., Souciet J.-L., Steensma H.Y., Talla E., Thierry A., Vandenbosch M., van der Hart Q.J.M., Van Dyk L., Vanoni M., Verhasselt P., Voet M., Volckaert G., Wambutt R., Watson M.D., Weber N., Wedler E., Wimpfli P., Wolf K., Wright J.F., Zaccaria P., Zimmermann M., Zoliner A., Kleine K.;		CC -!- SUBUNIT: Interacts with ADA2. Part of the ADA/GCN5 complex that consists of HFL1/ADA1, ADA2, ADA3, SPT0/ADAS and GCN5. Component of the SAGA complex, at least composed of SPT2, SPT7, SPT8, SPT20/ADAS, HFL1, ADA2, ADA3/NGFL, TRAI and GCN5.		
RX				CC -!- SIMILARITY: Contains 1 bromodomain.
RA				CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
RA				CC -!- SIMILARITY: Contains 1 bromodomain.
RA				CC EMBL; X68528; CAA44602.1;
RA				DR SGD; S0003484; GCN5
RA				DR EMBL; Z73037; CAA97281.1;
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				DR SGD; S0003484; GCN5
RA				DR GO; GO:0000124; C1SAGA complex; IDA.
RA				DR InterPro; IPR001487; Bromodomain.
RA				DR PIR; S28051; S28051.
RA				DR PDB; 1YGH; 02-AUG-91.
RA				DR PDB; 1E6I; 24-NOV-00.
RA				DR Germoline; J-14164;
RA				DR TRANSFAC; T02145;
RA				

PROSITE: PS00633; BRONODOMAIN\_1; 1.  
PROSITE: PS50014; BRONODOMAIN\_2; 1.  
KW: Transcription regulator; Transf erase; Activator; Chromatin regulator;  
KW: Trans-acting factor; Nuclear protein; Bromodomain; 3D-structure.  
ACT SITE: 173  
GENERAL DOMAIN: BRONODOMAIN.  
DORN: 344 414  
FT STRAND 173 414  
FT STRAND 100 105  
FT HELIX 111 127  
FT TURN 129 130  
FT HELIX 133 141  
FT TURN 143 144  
FT STRAND 146 152  
FT TURN 153 155  
FT STRAND 156 166  
FT HELIX 167 169  
FT TURN 170 170  
FT STRAND 171 179  
FT TURN 181 182  
FT TURN 186 187  
FT HELIX 188 203  
FT STRAND 208 213  
FT HELIX 215 217  
FT HELIX 218 223  
FT TURN 224 225  
FT STRAND 227 227  
FT HELIX 234 237  
FT STRAND 238 238  
FT TURN 239 239  
FT STRAND 241 241  
FT TURN 244 245  
FT STRAND 249 253  
SEQUENCE: 439 AA; 51069 MW; 320070DDCTBF70D CRC64;  
Query Match 30.1%; Score 182; DB 1; Length 439;  
Best Local Similarity 38.5%; Pred. No. 5.3e-12; Indels 2; Gaps 1;  
Matches 37; Conservative 16; Mismatches 41; RT  
QY 16 VMAQLWKGHEQAMPPROPYDAVKLGLPDYKTIKOPMDMGTTKERLENNYYWAASECMDF 75  
Db 339 1LTLEQNHAAMAPQDLYV-KEEVDYDFIKEPKMDLSTMEKLESNKYQRMEDFYDA 396  
QY 76 NTMFTNCYIYINKPDDIVMAQLEKIFKQYRASMP 111  
Db 397 RLVEFNCRRYNGENTSYYYANLUEKKFFNFKVKEIP 432  
RESULT 9  
P300\_HUMAN STANDARD PRT: 2414 AA.  
ID P300\_HUMAN STANDARD  
AC Q09172;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DE P300-associated protein p300 (EC 2.3.1.48).  
GN EP300 OR P300.  
OC Bivalvales (Human);  
OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9666;  
RN [1]  
SEQUENCE FROM N.A.  
MEDLINE=95011587; PubMed=753245;  
RA Eckner R., Ewen M.E., Newsome D., Gerdes M., Decaprio J.A.,  
RA Lawrence J.B., Livingston D.M.;  
RT "Molecular cloning and functional analysis of the adenovirus E1A-  
associated 300-kD protein (p300) reveals a protein with properties of  
a transcriptional adaptor.";  
RN Genes Dev. 8:869-884 (1994).  
RN [2]  
ENZYMATIC ACTIVITY: PubMed=8945521;  
RX MEDLINE=97100994; PubMed=8945521;  
RA Ogryzko V.V., Schultz R.L., Russianova V., Howard B.H., Nakatani Y.;  
RT "The transcriptional coactivators p300 and CBP are histone  
acetyltransferases.";  
RN [3]  
INTERACTION WITH PCAF.  
MEDLINE=96300317; PubMed=8684459;  
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.;  
RT "A P300/CBP-associated factor that competes with the adenoviral  
oncoprotein E1A.";  
RN Nature 382:319-324 (1996).  
[4]  
INTERACTION WITH HIF1A AND CREBBP.  
MEDLINE=97075102; PubMed=8917528;  
RX Arany Z., Huang L.E., Echner R., Bhattacharyya S., Jiang C.,  
RA Goldberg M.A., Bunn H.F., Livingston D.M.;  
RT "An essential role for p300/CBP in the cellular response to hypoxia.";  
Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973 (1996).  
[5]  
INTERACTIONS WITH NR3C1.  
MEDLINE=9830578; PubMed=9509696;  
RA Fryer C.J., Archer T.K.;  
RT "Chromatin remodelling by the glucocorticoid receptor requires the  
BRG1 complex.";  
RN Nature 393:88-91 (1998).  
[6]  
INTERACTION WITH NCOA6.  
MEDLINE=20383976; PubMed=10833961;  
RA Ko L., Cardona G.R., Chin W.W.;  
RA Kishi N., Matsuno K., Nakamura K., Weinmaster G., Okano H.,  
RA Kishimoto M.;  
RT "Role of Dlx2-1 as a transcriptional regulator downstream of the  
Thyroid receptor.";  
RN [7]  
INTERACTION WITH DDX1.  
MEDLINE=21577616; PubMed=11564735;  
RA Yamamoto N., Yamamoto S.-I., Inagaki F., Kawaiuchi M., Fukamizu A.,  
RA Kishimoto M.;  
RT "Identification of a promoter-specific transcriptional activation  
domain at the C-terminus of the Wnt-effector protein TCF4.";  
RN [8]  
PHOSPHORYLATION AND INTERACTION WITH TCF7L2 AND LEF1.  
MEDLINE=22450614; PubMed=12446687;  
RA Hecht A., Stemmler M.P.;  
RA Wagner G., Eck M.J.;  
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible  
factor-1 alpha.";  
RN [9]  
VARIANT BREAST CANCER PRO-827 AND GLY-1013, VARIANT PANCREATIC CANCER  
TYR-1650, AND VARIANT COLORECTAL CANCER GLN-2221.  
RP RT  
RT STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.  
RX MEDLINE=20164312; PubMed=10700188;  
RA Gayther S.A., Batley S.J., Liniger L., Bannister A., Thorpe K.,  
RA Chin S.-F., Daggio Y., Russell P., Wilson A., Sowter H.M.,  
RA Delhatty J.D.A., Ponder B.A.J., Kouzarides T., Caldas C.;  
RT "Mutations truncating the EP300 acetylase in human cancers.";  
RN Nat. Genet. 24:300-303 (2000).  
RN [10]  
STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.  
RP RT  
RX MEDLINE=21957234; PubMed=11959990;  
RA Freedman S.J., Sun J., Poy F., Kung A.L., Livingston D.M.,  
RA Wagner G., Eck M.J.;  
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible  
factor-1 alpha.";  
RN [11]  
STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.  
RP RT  
RX MEDLINE=21957234; PubMed=11959990;  
RA Freedman S.J., Sun J., Poy F., Kung A.L., Livingston D.M.,  
RA Wagner G., Eck M.J.;  
RT "Structural basis for recruitment of the adenovirus E1A Protein.  
Proc. Natl. Acad. Sci. U.S.A. 99:5367-5372 (2002).  
CC CC  
FUNCTION: Functions as histone acetyltransferase and regulates  
transcription via chromatin remodeling. Acetylates all four core  
histones in nucleosomes. Histone acetylation gives an epigenetic  
tag for transcriptional activation. Binds to and may be involved  
in the transforming capacity of the adenovirus E1A Protein.  
CC CC  
CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.  
CC CC  
SUBUNIT: Interacts with NR3C1, PCAF, NCOA6, LBF1 and TCF7L2. The  
TAF-type 1 domain interacts with HIF1A. Probably part of a complex  
CC CC  
with HIF1A and CREBBP. Interacts with Dix1.  
CC CC  
SUBCELLULAR LOCATION: Nuclear.

CC -!- PTM: Phosphorylated.  
 CC -!- DEGRADE: Detects in EP300 may play a role in epithelial cancer.  
 CC -!- SIMILARITY: Contains 1 bromodomain.  
 CC -!- SIMILARITY: Contains 1 ZZ-type zinc finger.  
 CC -!- SIMILARITY: Contains 2 TAZ-type zinc fingers.  
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/P300ID97.html".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires an agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC

CC EMBL: U01877; AAA18639.1; - .  
 DR PIR: A54277; A54277.  
 DR IL3E; 11-PEB-03.  
 DR TRANSFAC; T01427; - .  
 DR Genew; HGNC; 3373; EP300.  
 DR MIM; 602700; - .  
 DR GO: GO:0008022; E:protein C-terminus binding; TAS.  
 DR GO; GO:003713; F:transcription co-activator activity; TAS.  
 DR GO; GO:003700; F:transcription factor activity; TAS.  
 DR GO; GO:007399; F:neurogenesis; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR001487; Bromodomain.  
 DR InterPro; IPR00101; KIX.  
 DR InterPro; IPR000433; Znf\_TAZ.  
 DR Pfam; PF00439; bromodomain; 1.  
 DR Pfam; PF02172; KIX; 1.  
 DR Pfam; PF02135; ZF\_TAZ; 2.  
 DR Pfam; PF00559; ZZ; 1.  
 DR SMART; SM00297; BRONO; 1.  
 DR SMART; SM0051; Znf\_TAZ; 2.  
 DR PROSITE; PS00653; BROMODOMAIN\_1; 1.  
 DR PROSITE; PS00054; BROMODOMAIN\_2; 1.  
 DR PROSITE; PS0134; ZF\_TAZ; 2.  
 DR PROSITE; PS01357; ZF\_ZZ; 1.  
 DR PROSITE; PS00135; ZF\_ZZ; 1.  
 DR Transcription; Transcription regulation; Nuclear protein; Bromodomain;  
 KW Cell cycle; Zinc-finger; Phosphorylation; Disease mutation;  
 KW 3D-structure.  
 FT DOMAIN 11 17 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT ZN\_FING 331 417 TAZ-TYPE 1.  
 FT DOMAIN 1067 1139 BROMODOMAIN.  
 FT ZN\_FING 1572 1818 BINDING REGION FOR E1A ADENOVIRUS.  
 FT ZN\_FING 1664 1707 ZZ-TYPE.  
 FT DOMAIN 1728 1709 TAZ-TYPE 2.  
 FT DOMAIN 797 800 POLY-SER.  
 FT DOMAIN 1519 1516 POLY-GLU.  
 FT DOMAIN 2066 2059 POLY-GLN.  
 FT VARIANT 827 827 L->P (in breast cancer).  
 FT VARIANT 1013 1013 /FTID=VAR\_014428.  
 FT VARIANT 1650 1650 E->G (in breast cancer).  
 FT VARIANT 2221 2221 S->Y (in pancreatic cancer).  
 FT SEQUENCE 2414 AA; 264143 MW; 6BFF909E849D693 CRC64;

SQ Query Match 30.0%; Score 1815; DB 1; Length 2414;  
 Best Local Similarity 38.0%; Pred. No. 3.6e-11;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

RA 17 MKALWKHQ-PFWAPFROPDVAKLGLPDTYHIIKQPMDMGTCIKRLENNYYAASECMDF 75  
 RP ACETYLATION OF NCOA3.  
 RA MEDLINE=99418638; PubMed=10490106;

Db 1062 LEALYRDPESLPRQVDPOLGLPFDIVKSPMDLSTIKRKLDTGQYQEPHQYVDDI 1121  
 QY 76 NMFTQIYIYKPTDDIVLMAQTLERKIFLQKV 107  
 Db 1122 WLMFNNAWLYNRKTSVYKYSKSLSEVFEQEI 1153  
 CC RESULT 10  
 CC CBP HUMAN STANDARD; PRT; 2442 AA.  
 CC ID CBP\_HUMAN STANDARD;  
 CC AC 092793; Q00147; Q16376;  
 CC DT 15-JUL-1998 (Ref. 36, Created)  
 CC DT 15-JUL-1998 (Ref. 36, Last sequence update)  
 CC DT 10-OCT-2003 (Ref. 42, Last annotation update)  
 CC DE CREBB-binding protein (EC 2.3.1.48).  
 CC GN CREBBP OR CBP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catazrhini; Hominidae; Homo.  
 NCBI\_TAXID=9606  
 RN RN SEQUENCE FROM N. A.  
 RX MEDLINE=97385172; PubMed=9238046;  
 RA Sobulo O.M., Borrow J., Tomek R., Reshimi S., Harden A.,  
 RA Schlegelberger B., Housman D., Doggett N.A., Rowley J.D.,  
 RA Zeleznik-Le N.J.,  
 RA "ML is fused to CBP, a histone acetyltransferase, in therapy-related  
 RT acute myeloid leukemia with a t(11;16)(q23;p13.3).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:732-837(1997).  
 RN [1]  
 RN SEQUENCE FROM N. A.  
 RX MEDLINE=9721049; PubMed=9177780;  
 RA Giles R.H., Petri F., Dauwe R.H.G., den Hollander A.I.,  
 RA Lushnikova T., van Ommer G.J.B., Goodman R.H., Deaven L.L.,  
 RA Doggett N.A., Peters D.J.M., Breuning M.H.,  
 RA "Construction of a 1.2-Mb contig surrounding, and molecular analysis  
 RT of, the human CREB-binding protein (CBP/CREBBP) gene on chromosome  
 RT 16p13.3.";  
 RL Genomics 42:96-144 (1997).  
 RN [2]  
 RN SEQUENCE OF 1-405 FROM N. A.  
 RX MEDLINE=96316966; PubMed=8782817;  
 RA Borrow J., Stanton V.P., Andriesen J.M., Becher R., Behm F.G.,  
 RA Chagastis R.S.K., Civin C.I., Distefano C., Dube I., Frischkrauf A.M.,  
 RA Horsman D., Mirelman E., Volini S., Witmore R.,  
 RA "A p300/CBP-associated factor that competes with the adenoviral  
 RT oncoprotein E1A";  
 RL Nature 382:319-324 (1996).  
 RN [3]  
 RN SEQUENCE OF 1-405 FROM N. A.  
 RX MEDLINE=96316966; PubMed=8782817;  
 RA Borrow J., Stanton V.P., Andriesen J.M., Becher R., Behm F.G.,  
 RA Chagastis R.S.K., Civin C.I., Distefano C., Dube I., Frischkrauf A.M.,  
 RA Horsman D., Mirelman E., Volini S., Witmore R.,  
 RA "The transitive regulation t(8;16)(p11;p13) of acute myeloid leukemia fuses  
 RT a putative acetyltransferase to the CREB-binding protein.";  
 RL Nature 382:319-324 (1996).  
 RN [4]  
 RN INTERACTION WITH PCAF.  
 RX MEDLINE=96300317; PubMed=8684459;  
 RA Yang X.-J., Oryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.,  
 RA Goldberg M.A., Bunn H.F., Livingston D.M.,  
 RA Arany Z., Huang L.E., Echner R., Shattuckay S., Jiang C.,  
 RA "An essential role for p100/CBP in the cellular response to hypoxia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973 (1996).  
 RN [5]  
 RN INTERACTION WITH PML.  
 RX MEDLINE=99178939; PubMed=10077561;  
 RA Doucas V., Tini M., Egan D.A., Evans R.M.,  
 RA "Modulation of CREB binding protein function by the promyelocytic  
 RT (PML) oncoprotein suggests a role for nuclear bodies in hormone  
 RT signaling.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2627-2632 (1999).  
 RN [6]  
 RN INTERACTION WITH PML.  
 RX MEDLINE=99178939; PubMed=10077561;  
 RA Doucas V., Tini M., Egan D.A., Evans R.M.,  
 RA "Modulation of CREB binding protein function by the promyelocytic  
 RT (PML) oncoprotein suggests a role for nuclear bodies in hormone  
 RT signaling.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2627-2632 (1999).  
 RN [7]  
 RN ACETYLATION OF NCOA3.  
 RX MEDLINE=99418638; PubMed=10490106;

RA	Chen H., Lin R.J., Xie W., Wilpitz D., Evans R.M.;	DR	EMBL; U89354; AAC51339.1;
RT	"Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase.";	DR	EMBL; U89355; AAC51340.1;
RT	"Activation of an acetylase.";	DR	PDB; 1LBC; 24-APR-02.
RL	"Cell 98:675-686 (1999).";	DR	PDB; 1JSP; 17-AUG-02.
RN	"[8] INTERACTION WITH NCOA6.";	DR	TRANSFAC; T02234;
RP	"MEDLINE=2032329; PubMed=108666662;";	DR	DR GenBank; HGNC; 2348; CREBBP.
RA	"Mahajan M.A., Samuels H.H.;	DR	DR MIM; 600140;
RT	"A new family of nuclear receptor co-regulators that integrates nuclear receptor signaling through CBP.";	DR	DR MIM; 180819;
RL	"Mol. Cell. Biol. 20:5048-5063 (2000).";	DR	DR GO; GO:0005737; C:cytoplasm; TAS.
RN	"[9] IDENTIFICATION IN A COMPLEX WITH NCOA2; NCOA3; IKKA; IKKB AND IKBGK. MEDLINE=21968197; PubMed=11971985;";	DR	DR GO; GO:0005634; C:nucleus; TAS.
RA	"Dames S.A., Martinez-Yamout M., De Guzman R.N., Dyson H.J., Wright P.E.;	DR	DR GO; GO:0003713; F:transcription co-activator activity; TAS.
RA	"Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response.";	DR	DR GO; GO:0003700; F:transcription factor activity; TAS.
RT	"RT PRO-1378.";	DR	DR GO; GO:0006461; P:protein complex assembly; TAS.
RT	"Regulation of SRC-3 (pCIP/ACTR/AIB-1/RAC-3/TRAM-1) coactivator activity by I kappa B kinase.";	DR	DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; TAS.
RL	"Mol. Cell. Biol. 22:3549-3561 (2002).";	DR	DR InterPro; IPR001497; Bromodomain.
RN	"[10] STRUCTURE BY NMR OF CP 345-439 IN COMPLEX WITH 776-826 OF HIF1A. MEDLINE=21957241; PubMed=11959777;";	DR	DR InterPro; IPR000197; TAZ finger.
RA	"Dames S.A., Martinez-Yamout M., De Guzman R.N., Dyson H.J., Wright P.E.;	DR	DR IPRO00433; Znf_ZZ.
RA	"Wu R.C., Qin J., Hashimoto Y., Wong J., Xu J., Tsai S.Y., Tsai M.J., O'Malley B.W.;	DR	DR PF00439; bromodomain; 1.
RT	"Regulation of SRC-3 (pCIP/ACTR/AIB-1/RAC-3/TRAM-1) coactivator activity by I kappa B kinase.";	DR	DR PF02172; KIX; 1.
RT	"Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response.";	DR	DR PF02135; zf-TAZ; 2.
RL	"Proc. Natl. Acad. Sci. U.S.A. 99:5271-5276 (2002).";	DR	DR PF00569; ZZ; 1.
RN	"[11] VARIANT RTTS PRO-1378.";	DR	DR PRINTS; PR00563; BRMOMODomain.
RP	"MEDLINE=21231140; PubMed=11331617;";	DR	DR SMART; SM00297; BRMOMODomain.
RA	"Murata T., Kurokawa R., Krone A., Tatsutani K., Ishii M., Taki T., Masuno M., Ohashi H., Yanagisawa M., Rosenfeld M.G., Glass C.K., Hayashi Y.;	DR	DR SM00551; Znf_TAZ; 1.
RT	"Defect of histone acetyltransferase activity of the nuclear transcriptional coactivator CBP in Rubinstein-Taybi syndrome.";	DR	DR PROSITE; PS50014; BRMOMODomain_2; 1.
RT	"Transcriptional coactivator 107R in 2001).";	DR	DR PROSITE; PS50134; ZF-TAZ_2; 2.
CC	"- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a specific tag for transcriptional activation. Also acetylates non-histone proteins, like NCOA3 coactivator. Mediates cAMP-gene regulation by binding specifically to phosphorylated CREB protein. CBP, is coactivator, augments the activity of phosphorylated CREB to activate transcription of cAMP-response genes.";	DR	DR PROSITE; PS50137; ZF_ZZ; 1.
CC	"- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.";	DR	DR PROSITE; PS50135; ZF_ZZ; 1.
CC	"- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, pCAF and PML. The TAZ type 1 domain interacts with HIF1A. Found in a complex containing NCOA2; NCOA3; IKKA; IKKB and IKBGK. Interacts with EP300 coactivator. Probably part of a complex with HIF1a and EP300.";	DR	DR KW Bromodomain; Chromosomal translocation; Nuclear protein; Activator; Disease mutation; Zinc-finger; Repeat; Disease mutation; 3D-structure.
CC	"- SUBCELLULAR LOCATION: Nuclear.";	DR	DR KW Disease mutation; ZN_FING; 1701; 1744.
CC	"- DISORDER: Involved in acute myeloid leukemias through chromosomal translocations t(8;16)(p11;p13) involving MYSR3 and CREBBP, and t(11;16)(q23;p13.3) involving MLL/HRX and CREBBP.";	DR	DR FT_ZN_FING; 347; 433.
CC	"- DISORDRE: Defects in CREBBP are the cause of Rubinstein-Taybi syndrome (RTTS) (NM_180849). RTTS is an autosomal dominant disorder characterized by craniofacial abnormalities, broad thumbs, broad big toes, mental retardation and a propensity for development of malignancies.";	DR	DR FT_DOMAIN; 452; 683.
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FT_DOMAIN; 1103; 1175.
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR FT_DOMAIN; 1061; 1064.
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR FT_DOMAIN; 1199; 1487.
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR FT_DOMAIN; 1555; 1562.
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FT_ZN_FING; 1755; 1846.
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR FT_DOMAIN; 1943; 1948.
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR FT_DOMAIN; 1967; 1970.
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR FT_DOMAIN; 2081; 2085.
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FT_DOMAIN; 2199; 2216.
CC	"- SUBCELLULAR LOCATION: Nuclear.";	DR	DR FT_DOMAIN; 2245; 2248.
CC	"- DISORDER: Involved in acute myeloid leukemias through chromosomal translocations t(8;16)(p11;p13) involving MYSR3 and CREBBP, and t(11;16)(q23;p13.3) involving MLL/HRX and CREBBP.";	DR	DR FT_DOMAIN; 2297; 2300.
CC	"- DISORDRE: Defects in CREBBP are the cause of Rubinstein-Taybi syndrome (RTTS) (NM_180849). RTTS is an autosomal dominant disorder characterized by craniofacial abnormalities, broad thumbs, broad big toes, mental retardation and a propensity for development of malignancies.";	DR	DR FT_VARIANT; 1378; 1378.
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FAE > NSG (IN REF. 2).
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR FAE > VAR_015578.
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR FAE > P (IN REF. 2).
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR FAE > R (IN REF. 2).
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FAE > VV (IN REF. 2).
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR FAE > L (IN REF. 2).
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR FAE > N (IN REF. 2).
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR FAE > P (IN REF. 2).
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR FAE > T (IN REF. 2).
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR FAE > AA; 265336 MW; 42D084619475F3D2 CRC64;
CC	"- Best Local Similarity 35%; Conservative 20%; Mismatches 36%; Indels 1; Gaps 1; Score 180.5; DB 1; Length 2442;	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 1 bromodomain.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- SIMILARITY: Contains 2 TAZ-type zinc fingers.";	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; www=http://www.intobioigen.fr/services/chromcancer/Genes/CBPPD42.html".	DR	DR Score 180.5; DB 1; Length 2442;
CC	"- This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).";	DR	DR

RESULT 11  
 BDF1\_YEAST\_ID -BDF1\_YEAST STANDARD; PRT; 686 AA.  
 AC P35817; Q06048;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE BDF1 Protein.  
 GN BDF1 OR YLR399C OR L8084 18.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukarya; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=5116323; PubMed=7816623;  
 RA Lygerou Z., Conesa C., Lesage P., Swanson R.N., Ruet A., Carlson M.,  
 RA Sentenac A., Seraphin B.;  
 RA "The yeast BDF1 gene encodes a transcription factor involved in the  
 expression of a broad class of genes including snRNAs.";  
 RL Nucleic Acids Res. 22:5332-5340(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Roeder G.S., Chua P.R.; Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=9131267; PubMed=9169871;  
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,  
 RA Benes V., Brueckner M., Della H., Dubois E., Duesterhoeft A.,  
 RA Enten K.-D., Fleisch M., Goffeau A., Hebling U., Heumann K.,  
 RA Heuss-Neitzel D., Hilbert H., Hilger P., Kleine K., Koetter P.,  
 RA Louis E.J., Messenguy F., Meves H.-W., Mioaga T., Mostl D., Pohl T.M.,  
 RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,  
 RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Kirke M., Rose M.,  
 RA Scherzer M., Scherzer S., Scholler P., Schwager C., Schwarz S.,  
 RA Underwood A.P., Urestarazu I.A., Vandemboul M., Verhasselt P.,  
 RA Vierendels F., Voet M., Volkstaert G., Voss H., Wambutt R., Wedler E.,  
 RA Wedler H., Zimmermann F.K., Zolinter A., Hahn J., Hohnsel J.D.;  
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIII.";  
 RL Nature 387:87-90(1997).  
 RN SEQUENCE OF 471-686 FROM N.A.  
 RX MEDLINE=91309447; PubMed=832235;  
 RA Widner W.R., Wickner R.B.;  
 RT "Evidence that the SK1 antiviral system of *Saccharomyces cerevisiae*  
 acts by blocking expression of viral mRNA.";  
 RL Mol. Cell. Biol. 13:4331-4341(1993).  
 CC -- FUNCTION: Transcription factor involved in the expression of a  
 broad class of genes including snRNAs. Required for sporulation.  
 CC -- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -- SIMILARITY: Contains 2 bromodomains.  
 CC -- SIMILARITY: Contains 1 ER domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC EMBL; Z1B944; CAA79377.1; -.  
 DR EMBL; U18116; AAA89115.1; -.  
 DR EMBL; U19729; AAA82357.1; -.  
 DR EMBL; L13469; AAA35048.1; -.  
 DR PIR; S55955; S55955. -.  
 DR GermOnline; 142462; -.  
 DR TRANSPAC; T03204; -.  
 DR SGD; S0004391; BDF1.

DR GO; GO:0005634; C: nucleus; IDA.  
 DR GO; GO:0030528; P: transcription regulator activity; IMP.  
 DR GO; GO:0007151; P: porcupine (sensu Saccharomyces); IMP.  
 DR InterPro; IPR001487; Bromodomain.  
 DR Pfam; PF00439; Bromodomain; 2.  
 DR PRINTS; PR00503; Bromodomain.  
 DR SMART; SM00297; BRON; 2.  
 DR PROSITE; PS000633; BROMODOMAIN 1; 2.  
 DR PROSITE; PS50014; BROMODOMAIN 2; 2.  
 DR Transcription regulation; Nuclear protein; Bromodomain; Repeat;  
 DR Sporulation.  
 FT DOMAIN 165 237 BROMODOMAIN 1.  
 FT DOMAIN 332 404 BROMODOMAIN 2.  
 FT DOMAIN 522 686 ET DOMAIN.  
 FT CONFLICT 8 8 Q -> LC (IN REF. 1).  
 FT CONFLICT 93 94 GA -> R (IN REF. 2).  
 FT CONFLICT 94 94 A -> P (IN REF. 2).  
 FT CONFLICT 282 282 A -> P (IN REF. 1).  
 FT CONFLICT 385 385 D -> E (IN REF. 1).  
 FT CONFLICT 493 493 A -> R (IN REF. 4).  
 SQ SEQUENCE 686 AA; BCCD52F41F91D0DA CRC64;  
 SQ SEQUENCE 76978 MW; BCCD52F41F91D0DA CRC64;

Query Match 29.1%; Score 176; DB 1; Length 686;  
 Best Local Similarity 31.0%; Pred. No. 3-7e-11;  
 Matches 35; Conservative 25; Mismatches 51; Indels 2; Gaps 1;

Qy 1 KPGRVINQOLQYLHKVVKMKAQWQ-KQFANPFRQPVDAVQKGLPDYHKLIKQPMMDGTTK 58  
 D5 310 KSKRLOQAKMFCQSYVQKELMAKGASVSYNNPFLEPVDPVSMNLPPYFDVYKPEPDGLTAK 369

RESULT 12  
 CES2\_HUMAN PRT; 1484 AA.  
 ID CES2\_HUMAN STANDARD;  
 AC Q96P9; Q9CC03;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cat eye syndrome critical region protein 2.  
 GN CECR2 OR XIAA1740.  
 OC Homo sapiens (Human).  
 OC Eukaryota; Metozoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_TaxID=9605;  
 RN [1] SEQUENCE FROM N.A. (ISOFORM A).  
 RX MEDLINE=21275466; PubMed=1381032;  
 RA Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riazi M.A.,  
 RA Bridgeland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,  
 RA Nguyen T., Fang F., Fu Y., Ray L., Wu H., Shauli S., Phan S., Yac Z.,  
 RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,  
 RA McDermid H.E.;  
 RT "Analysis of the cat eye syndrome critical region in humans and the  
 region of conserved synteny in mice: a search for candidate genes at  
 or near the human chromosome 22 pericentromere.";  
 RN Genome Res. 11:1053-1070(2001).  
 RN SEQUENCE OF 346-1484 FROM N.A. (ISOFORM A).  
 RC TISSUE-Brain;  
 RX MEDLINE=21082932; PubMed=1214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 for large proteins in vitro.";  
 RN DNA Res. 7:347-355(2000).  
 RN SEQUENCE OF 168-1484 FROM N.A. (ISOFORM B), AND INTERACTION WITH  
 RP LRPPRC.  
 RC TISSUE-Liver;



DR InterPro: IPR001965; Znf PHD.  
 DR Pfam: PF00439; bromodomain; 1.  
 DR Pfam: PF02172; KIX; 1.  
 DR Pfam: PF01335; zf-TAZ; 2.  
 DR Pfam: PF00569; ZZ; 1.  
 DR SMART; SN00297; BRM0; 1.  
 DR SMART; SN00551; Znf-TAZ; 2.  
 DR PROSITE; PS00033; BRM0DOMAIN; 1.  
 DR PROSITE; PS00014; BRM0DOMAIN; 2; 1.  
 DR PROSITE; PS0159; ZF-PHd; 1.  
 DR PROSITE; PS0134; ZF-TAZ; 2.  
 DR PROSITE; PS0137; ZF-ZZ; 1.  
 DR PROSITE; PS00135; ZF-ZZ; 2; 1.  
 DR Bromodomain; Metal-binding; zinc; Zinc-finger; Alternative splicing;  
 KW Repeat.  
 FT ZN\_FING 399 505 TAZ-TYPE 1.  
 FT DOMAIN 881 953 BRM0DOMAIN.  
 FT ZN\_FING 1493 1534 2Z-TYPE.  
 FT DOMAIN 1550 1631 TAZ-TYPE 2.  
 FT DOMAIN 1687 2008 GLY/GLN-RICH.  
 FT VARSPLIC 467 478 SDTTTGTTCRGSV -> F (in isoform a). /FTId=vSP\_000557  
 SQ SEQUENCE 2056 AA; 22179 MW; 949FF746086534P01 CRC64;  
 Query Match 26.8% Score 162; DB 1; Length 2056;  
 Best Local Similarity 35.9%; Pred. No. 3.7e-09;  
 Matches 33; Conservative 19; Mismatches 40; Indels 0; Gaps 0;  
 Qy 16 VMKALWKHQAPWAPFOPVAVKLGPDYKIIKTOPMDTIGTKEERLENNYYWASECMQDF 75  
 Db 876 VWEFLDKSEAAFPVDPVDDKLRLNPKDLYQYQVQAGQFCDDI 935  
 Qy 76 NTMFNCCYINNKPTDIVMAQTLKEIFLQKV 107  
 Db 936 WMLDNAWL1NRKNSKVYKVGKLUSEMVFSEM 967

RESULT 14

ID TFI1\_MOUSE STANDARD; PRT; 1051 AA.  
 AC Q64127; Q64126;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DE Transcription intermediary factor 1-alpha (TIF1-alpha) (Tripartite motif protein 24).  
 DE TIF1 OR TRIM4 OR TIF1A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI\_TaxID=10090;  
 OX [1]  
 RN SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC TISSUE=Carcinoma; PMID=7744009;  
 RX MEDLINE=9238825; PubMed=1247932;  
 RA le Douarin B.; Zechel C.; Garnier J.-M.; Lutz Y.; Tora L.; Pierrat B.; Heery D.; Gronemeyer H.; Chabon P.; Lissin R.;  
 RT "The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein Tr18." [2]  
 RT SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=9326626; PubMed=1247932;  
 RA Straussberg R.L.; Feingold E.H.; Grouse L.H.; Derge J.G.; Klausner R.D.; Collier F.S.; Wagner L.; Shenmen C.M.; Schulter G.D.; Altschul S.F.; Zeeberg B.; Butow K.H.; Schaefer C.F.; Bhat N.K.; Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.; Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;

DR	PRINTS; PRO00503; BROMODOMAIN.
DR	SMART; SM00512; BBC; 1.
DR	SMART; SM00316; BBOX; 2.
DR	SMART; SM00297; BROMO; 1.
DR	SMART; SM00249; BHD; 2.
DR	SMART; SM00184; RING; 1.
DR	PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR	PROSITE; PS50114; BROMODOMAIN_2; 1.
DR	PROSITE; PS50119; ZF_BBOX; 2.
DR	PROSITE; PS00359; ZF_PHD_1; 1.
DR	PROSITE; PS50016; ZF_PHD_2; 1.
DR	PROSITE; PS00518; ZF_RING_1; 1.
DR	PROSITE; PS00089; ZF_RING_2; 1.
DR	Transcription regulation; Repressor; DNA-binding; Bromodomain; Zinc-finger; Alternative splicing; Nuclear Protein; Coiled coil; Zinc-finger; Alternative splicing; Nuclear Protein; Coiled coil; Zinc-finger; Chromosomal translocation.
KW	Repeat; Proto-oncogene; Chromosomal translocation.
FT	DOMAIN 8 15 POLY-ALA.
FT	DOMAIN 19 22 POLY-ALA.
FT	ZN_FING 52 77 RING-TYPE.
FT	ZN_FING 158 211 B-BOX-TYPE 1.
FT	ZN_FING 218 259 B-BOX-TYPE 2.
FT	DOMAIN 289 359 COILED COIL (POTENTIAL).
FT	DOMAIN 344 347 POLY-GLIN.
FT	DOMAIN 583 587 POLY-SER.
FT	DOMAIN 755 780 NUCLEAR RECEPTOR BINDING SITE (NRBS).
FT	ZN_FING 827 874 PHD-TYPE.
FT	DOMAIN 892 908 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN 933 988 BROMODOMAIN.
FT	SITE 332 333 BREAKPOINT FOR TRANSLOCATION TO FORM TIF1A-BAF1 ONCOGENE.
FT	VARSPLIC 477 510 Missing (in isoform Short).
FT	SEQUENCE 1051 AA; 116656 MW; 610884C1C6885072 CRC64; /FID=VP_005773.
SQ	Query Match 26.5% Score 160.5; DB 1; Length 1051; Best Local Similarity 38.4%; Prd. No. 2.6e-09; Matches 33; Conservative 17; Mismatches 31; Indels 5; Gaps 2;
Qy	20 LANKHQPAWPERQPVDAVKGGLPDYHKIKQPMMDGTTIKRLENNY--YWAASBCMQDENT 77
Db	916 LYCHEMSTLAFQDPD--VPLTVPPYKXKTPNPMOLST-KKRLQEDYCHYTKPEDFVAEFLR 972
Qy	78 MFTNCYLYNKPTDIVIAQTLKIF 103
Db	973 IFQNCAAEFPDSEVANQIKLESYF 998
RESULT 15	
TP1A_HUMAN STANDARD; ID: O15165; C95854; AA: 1050 AA.	
[1] SEQUENCE FROM N.A. (ISOFORM SHORT).	
RC TISSUE-Breast cancer; MEDLINE=9115274; Therot S., Henrict C., Richefort H., Cavailles V.; Homo sapiens (Human).	
RCA Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Homidae; Homo. NCBI_TaxID=9606;	
RN [1] SEQUENCE FROM N.A. (ISOFORM SHORT).	
RP TISSUE-Breast cancer; MEDLINE=9127752; Pubmed=9115274; Therot S., Henrict C., Richefort H., Cavailles V.; Homo sapiens (Human).	
RPA Biot. Chem. 272:12062-12068 (1997); SEQUENCE FROM N.A. (ISOFORM SHORT).	
RN [2] MEDLINE=9914725; Pubmed=1002227; Venturini L., You J., Staelin M., Galien R., Lallemand V.	
RA De The H.; "TFIIGamma, a novel member of the transcriptional intermediary factor 1 family." Oncogene 18:1209-1217 (1999).	
RA	RA "TFIIGamma, a novel member of the transcriptional intermediary factor 1 family." Oncogene 18:1209-1217 (1999).
RT	RT SEQUENCE OF 477-510 (ISOFORM LONG).
RC	RC TISSUE=breast cancer; Cavailles V.; Submitted (JAN-1999) to Swiss-Prot.
RL	RL [4] INTERACTION WITH NR3C2.
RN	RN RP MELINE=20388588; Pubmed=10935545; Heilai-Levy C., Fagart J., Souque A., Wurtz J.-M., Moras D., Raesten-Oblin M.-E.; "Crucial role of the H11-H12 loop in stabilizing the active conformation of the human mineralocorticoid receptor." J. Endocrinol. 14:1210-1221 (2000).
RT	RT CC -1- FUNCTION: Interacts selectively in vitro with the AF2-activating domain of the estrogen receptors. Association with DNA-bound estrogen receptors requires the presence of estradiol.
RT	RT CC -1- SUBUNIT: Interacts with CBX1 and CBX3 (By similarity). Interacts with NR3C2.
RL	RL CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
RP	RP CC -1- ALTERNATIVE PRODUCTS:
RA	RA Event=Alternative splicing; Named isoforms=2; Name=Short; IsoId=015164-1; Sequence=Displayed; Name=Long; IsoId=015164-2; Sequence=VP_005772;
RA	RA CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
RA	RA CC -1- SIMILARITY: Contains 2 B box-type zinc fingers.
RA	RA CC -1- SIMILARITY: Contains 1 bromodomain.
RA	RA CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
RA	RA CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcements">http://www.isb-sib.ch/announcements</a> or send an email to license@isb-sib.ch).
RA	RA CC -1- DR EMPL: AF009333; AAB63385.1; -.
RA	RA DR EMPL; AP119022; AAD17258.1; -.
RA	RA DR HSSP; P29590; 1B0R.
RA	RA DR T2TRANSFAC; TO9495; -.
RA	RA DR Genew; HGNC:11812; TIF1L.
RA	RA DR MIM; 603406; -.
RA	RA DR GO:0005634; C:nucleus; TAS.
RA	RA DR GO:0005102; F:receptor binding; TAS.
RA	RA DR GO:0003743; P:transcription co-activator activity; TAS.
RA	RA DR GO:0006366; P:transcription from Pol II promoter; TAS.
RA	RA DR InterPro; IPR003449; Bbox_C.
RA	RA DR InterPro; IPR01187; Bromodomain.
RA	RA DR InterPro; IPR00315; Znf_Bbox.
RA	RA DR InterPro; IPR001565; Znf_PHD.
RA	RA DR InterPro; IPR001841; Znf_ring.
RA	RA DR Pfam; PF00439; bromodomain_1.
RA	RA DR Pfam; PF00638; PHD; 1.
RA	RA DR Pfam; PF00642; zf-B_box; 2.
RA	RA DR Pfam; PF00094; zf-B_box; 1.
RA	RA DR PRINTS; PRO1406; BBOXNFINGER.
RA	RA DR PRINTS; PRO0503; BRMDOMAIN.
RA	RA DR SNRT; SM00522; BBC; 1.
RA	RA DR SNRT; SM00336; BBOX; 2.
RA	RA DR SNRT; SM00297; BROMO; 1.
RA	RA DR SNRT; SM00249; PHD; 1.
RA	RA DR PRINTS; PRO1184; RING; 1.
RA	RA DR PROSITE; PS00633; BRMDOMAIN_1; FALSE_NEG.
RA	RA DR PROSITE; PS50104; BRMDOMAIN_2; 1.
RA	RA DR PROSITE; PS00119; 2F_BB0X; 2.
RA	RA DR PROSITE; PS00156; 2F_BB0X; 1.
RA	RA DR PROSITE; PS00135; 2F_BB0X; 1.

Search completed: July 9, 2004, 13:06:42

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:03:30 ; Search time 39 Seconds  
 (without alignments)

Sequence: 898.013 Million cell updates/sec

Title: US-09-784-553C-19

Perfect score: 605

Sequence: 1 KPGRVTNQQLQYHLKVMKAL.....IVLMATQTLKIFLQKVASMP 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31551802 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL 25:\*

1: sp\_archea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rrodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteri:\*

17: sp\_archeap:\*

Ques66 mus musculu

Q86hff6 mus musculu

Q86v88 homo sapien

Q8bum2 oryzae lat

Q8irng drosophila

Q9w413 drosophila

Q9awx9 brachydanio

Q9gub61 myxine glut

Q7zv99 brachydanio

Q8t775 brachioisto

Q8lw16 homo sapien

Q14789 homo sapien

Q9zta6 homo sapien

Q9iy44 mus musculu

Q91686 xenopus lae

Q20947 caenorhabdi

Q8t326 drosophila

Q9vcg6 drosophila

Q86579 caenorhabdi

Q95y80 caenorhabdi

Q20948 caenorhabdi

Q57335 brachydanio

Q9vbs1 drosophila

Q8tiz8 drosophila

Q9iy6 arabidopsis

Q9iyaz arabidopsis

Q9yin0 schizosaccharomyces

Q9ik27 arabidopsis

Q84xv2 arabidopsis

## ALIGNMENTS

RESULT 1  
 Q99pc5  
 ID Q99pc5  
 AC Q99pc5;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DR 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Ring 3 (Fragment).  
 GN BRD2.  
 OS Mus musculus (Mouse).  
 OC Bukaaryote; Metazoa; Chordata; Craniata; Vertebrata; Buteleostoni;  
 OC Mammalia; Buthrotia; Rodentia; Sciurognathi; Murinae; Muridae; Murinae; Mus.  
 RN [1] - TaxID:10000;  
 RN [1] -  
 RP SEQUENCE FROM N.A.

RC STRA\_N=C57BL;

RA Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW; B42A566963981F28 CRC64;

SQ SEQUENCE 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR MGI: MGI:93495; Brd2.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM00297; BROMO; 2.

DR PROSITE: PS00633; BROMODomain\_1; 2.

DR PROSITE; PS50044; BROMODomain\_2; 2.

DR NON\_TIR 503 503 AA; 55523 MW;

Query Match 100.0%; Score 605; DB 11; Length 503;

Best Local Simililarity 100.0%; N mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0; Pred. No. 5e-59;

DR Korf I.;

RT "Complete sequence of UL26B06.";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR01487; Bromodomain.

DR PFAM: PF00439; bromodomain; 2.

DR PRINTS: PR0503; BROMODomain.

DR SMART: SM0029

Qy	61	ENYYWAASEMQDNTMFNCYIYNKPIDDIVMAQLEKIFQKVASMP	111	RP	SEQUENCE OF 47-549 FROM N.A.
Db	130	ENYYWAASEMQDNTMFNCYIYNKPIDDIVMAQLEKIFQKVASMP	180	RC	STRAIN=INBREED CD-1; TISSUE=Testis;
				RA	Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,
				RQ	Okumura K., Kimura M., Inoko H., Inoko H.,
				RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RESULT 2				RN	[4]
O88411		PRELIMINARY;	PRT;	RP	SEQUENCE FROM N.A.
ID	O88411;			RA	STRAIN=129SvJ;
AC	O88411;			RA	Rowen L., Qin S., Nadan A., Loretz C., James R., Dors M., Mix L.,
DT	01-NOV-1998	(TREMBREL. 08, Created)		RA	Hall J., Lasky S., Hood L.,
DT	01-NOV-1998	(TREMBREL. 08, Last sequence update)		RA	"Sequence of the mouse major histocompatibility locus Class II region."
DT	01-OCT-2002	(TREMBREL. 22, Last annotation update)		RT	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DE		Female sterile homeotic related protein Frg-1.		RL	
GN	BRD2.			DR	EMBL; AL009226; CAA15818.1;
OS	Mus musculus (Mouse).			DR	EMBL; AL009226; CAA15819.1;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			DR	EMBL; AB010248; BAA24379.1;
OX	NCBI_TaxID=10090;			DR	EMBL; AB010247; BAA24378.1;
RN				DR	EMBL; AB010246; BAA24377.1;
RP	SEQUENCE FROM N.A.			DR	EMBL; AF100956; AAC29907.1;
RA	Rhee K., Bruno I.W., Besset V., Wolgemuth D.J.,			DR	HSSP; Q92831; 1B91.
RT	"Expression and potential role of Frg-1, a putative murine bromodomain-containing homologue of the Drosophila gene female sterile homeotic",			DR	InterPro; IPR001487; Bromodomain.
RT	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.			DR	pfam; PF0043; bromodomain_2.
DR	EMBL; AF015462; AAC24810.1; -.			DR	PRINT; PR0050; BRONO; 2.
DR	HSSP; Q9B31; 1B91.			DR	SMART; SM00297; BRONO; 2.
DR	DRD; MGI: 93495; Bird; -.			DR	PROSITE; PS000633; BROMODOMAIN_1; 2.
DR	InterPro; IPR001487; Bromodomain.			DR	PROSITE; PS50014; BROMODOMAIN_2; 2.
DR	PFam: PF00439; bromodomain_2.			SQ	SEQUENCE 798 AA; 88066 MW; 08D57FBE1385E96 CRC64;
DR	PRINTS; PR00503; BROMODOMAIN_2.			Query Match	100.0%; Score 605; DB 11; Length 798;
DR	SMART; SM00297; BRONO; 2.			Best Local Similarity	100.0%; Pred. No. 8.4e-59;
DR	PROSITE; PS000633; BROMODOMAIN_1; 2.			Matches	111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR	PROSITE; PS50014; BROMODOMAIN_2; 2.			Qy	1 KPGRTVNQLQYLHKYTMKALWKHOPAWPFPQVDIVKGLPDYKIKIOPMDNGTICKRL 60
SQ	SEQUENCE 798 AA; 88063 MW; _A9942517CF15B7A1 CRC64;			Db	70 KPGRTVNQLQYLHKYTMKALWKHOPAWPFPQVDIVKGLPDYKIKIOPMDNGTICKRL 129
Qy	1 KPGRTVNQLQYLHKYTMKALWKHOPAWPFPQVDIVKGLPDYKIKIOPMDNGTICKRL 60			Qy	61 ENYYWAASEMQDNTMFNCYIYNKPIDDIVMAQLEKIFQKVASMP 111
Db	70 KPGRTVNQLQYLHKYTMKALWKHOPAWPFPQVDIVKGLPDYKIKIOPMDNGTICKRL 129			Db	130 ENYYWAASEMQDNTMFNCYIYNKPIDDIVMAQLEKIFQKVASMP 180
RESULT 4				RESULT 4	
O90971				ID	Q90971; PRELIMINARY;
				AC	Q90971; PRT; 729 AA.
				AC	Q90971; (TREMBREL. 01, Created)
				DT	01-NOV-1996 (TREMBREL. 01, Last sequence update)
				DT	01-NOV-1996 (TREMBREL. 01, Last annotation update)
				DN	RING3.
				OS	Gallus Gallus (Chicken).
				OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Osteichthyes; Teleostei; Aves; Galliformes; Phasianidae; Phasianidae;
				OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;
				OC	Gallus.
				OX	NCBI_TAXID=9031;
				RN	[1]
				RC	SEQUENCE FROM N.A.
				RC	STRAIN=B12;
				RC	MEDLINE=93376536; PubMed=8781126;
				RC	Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
				RC	"Phylogeny and structure of the RING3 gene.";
				RC	SMART; SM00297; BRONO; 2.
				RC	PROSITE; PS000633; BROMODOMAIN_1; 2.
				RC	PROSITE; PS50014; BROMODOMAIN_2; 2.
				SQ	SEQUENCE 729 AA; 80389 MW; 06B1B92C804DF7B8 CRC64;
				Q3	[3]
				RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
				RL	InterPro; IPR001487; Bromodomain.
				DR	pfam; PF00439; bromodomain_2.
				DR	PRINT; PR0050; BROMODOMAIN.
				DR	SMART; SM00297; BRONO; 2.
				DR	PROSITE; PS000633; BROMODOMAIN_1; 2.
				DR	PROSITE; PS50014; BROMODOMAIN_2; 2.
				SQ	SEQUENCE 729 AA; 80389 MW; 06B1B92C804DF7B8 CRC64;



RESULT 8	Q8N5M3	PRELIMINARY;	PRT;	556 AA.	Db	30 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 89
ID Q8N5M3	PRELIMINARY;	PRT;	556 AA.	Qy	61 ENNYWAASECMQDFNTMFTNCYLYNPKPTDDIVLMAQALEKIFLQKVASMP 111	
AC Q8N5M3 ;				Db	90 ENNYWAASECMQDFNTMFTNCYLYNPKPTDDIVLMAQALEKIFLQKVASMP 140	
DT 01-OCT-2002 (TREMBrel. 22, Created)						
DT 01-OCT-2002 (TREMBrel. 22, Last sequence update)						
DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)						
OS Homo sapiens (Human)						
OC Mammalia; Eutheria; Chordata; Primates; Catarrhini; Hominidae; Homo.						
OX NCBI_TaxID=9606;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Colon;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC012124; AAH32124; 1; -.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS00634; BROMODOMAIN_2; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 556 AA; 60942 MW; 8352F5DF1801A793 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 4; Length 556;						
Matches 97; Conservative 5; Pred. No. 6.1e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Testis;						
RA Shang E.; Wolgemuth D.J.;						
RA Bromodomain-containing homolog of the Drosophila gene female sterile homeotic",						
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.						
RL EMBL; AF263193; AAF78072.1; -.						
DR HSSP: Q92831; 1B91.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79745 MW; TAB3B4DAD38A78F4 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						
RA Strausberg R.;						
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.						
RL EMBL; BC01536; AAH31536.1; -.						
DR MGD; MGI:1914632; Brd3.						
DR InterPro: IPR001487; Bromodomain.						
DR Pfam: PF00439; bromodomain; 2.						
DR PRINTS: PR00503; BROMODDomain.						
DR SMART: SM00297; BROMO; 2.						
DR PROSITE: PS00633; BROMODOMAIN_1; 2.						
DR PROSITE: PS50014; BROMODOMAIN_2; 2.						
SQ SEQUENCE 726 AA; 79793 MW; 4BC2FCGACA41DEC0 CRC64;						
Query Match 1 KPGRKINQNLQMQNVVVKLWQFQAWPFQPVDIAKLNPDYKILKIQPMNDGTTKRL 60						
Best Local Similarity 89.6%; Score 542; DB 11; Length 726;						
Matches 97; Conservative 5; Pred. No. 8.2e-52; Indels 0; Gaps 0;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE=Body;						

RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RA	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RA	[2].
RA	SEQUENCE FROM N.A.
RA	RC TISSUE="Body";
RA	RA Strausberg R.;
RA	RA Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
RA	RL EMBL: BC055543; AAH55543.1; -.
RA	DR Hypothetical protein.
RA	KW FT
RA	NON TBR 515 515
RA	SEQUENCE 515 AA: 56253 MW: C3189F56736332C CRC64;
RA	Qy 1 KPGSVTNQLOQYIHKVVMRALKHOFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 60
RA	Db 27 KPGRKTNQLOQYMQNTVKTILWKGQFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 86
RA	Qy 61 ENYYWAASECMQDFNTMFTNCYIYNKPTDDIVMACTLEKIFLQVAVSMP 111
RA	Db 87 ESYVYASASCMQDFNTMFTNCYIYNKPTDDIVMAQALEKIFLQVAVLMP 137
RA	SEQUENCE FROM N.A.
RA	RC TISSUE="Body";
RA	RA Strausberg R.;
RA	RA Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
RA	RL EMBL: BC055533; AAH55533.1; -.
RA	DR Hypothetical protein.
RA	KW FT
RA	NON TBR 499 499
RA	SEQUENCE 499 AA: 55075 MW: 801BBDD64F9E7216 CRC64;
RA	Qy 1 KPGRVTNQLOQYIHKVVMRALKHOFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 60
RA	Db 27 KPGRKTNQLOQYMQNTVKTILWKGQFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 86
RA	Qy 61 ENYYWAASECMQDFNTMFTNCYIYNKPTDDIVMACTLEKIFLQVAVSMP 111
RA	Db 87 ENYYWAASECMQDFNTMFTNCYIYNKPTDDIVMAQALEKIFLQVAVSMP 137
RA	SEQUENCE FROM N.A.
RA	RC TISSUE="Body";
RA	RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.; Klausner D.D.; Colins F.S.; Wagner L.; Shemmer G.M.; Schuller G.D.,
RA	RA Zeiberger B.; Buetow K.H.; Schaefer C.F.; Bhatt N.K.,
RA	RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.,
RA	RA Ditchkoff L.; Marcusina K.; Farmer A.; Rubin G.M.; Hong L.,
RA	RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.,
RA	RA Raha S.S.; Loqueland N.A.; Kettman J.W.; Abramson R.D.; Mullahy S.J.,
RA	RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.,
RA	RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.,
RA	RA Villalobos D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.B.,
RA	RA Paine J.; Helton E.; Kettman M.; Madan A.; Rodrigues S.; Sanchez A.,
RA	RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.,
RA	RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.,
RA	RA Krywinski M.I.; Skalka U.; Smailus D.E.; Schnerr A.; Schein J.E.,
RA	RA Jones S.J.; Marra M.A.,
RA	SEQUENCE FROM N.A.
RA	RC TISSUE="Body";
RA	RA Strausberg R.;
RA	RA Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
RA	RL EMBL: BC055533; AAH55533.1; -.
RA	DR Hypothetical protein.
RA	KW FT
RA	NON TBR 515 515
RA	SEQUENCE 515 AA: 56253 MW: C3189F56736332C CRC64;
RA	Qy 1 KPGSVTNQLOQYIHKVVMRALKHOFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 60
RA	Db 27 KPGRKTNQLOQYMQNTVKTILWKGQFAWPFPQPVDAVKGFLPDYHKIKOPMDMGTKRRL 86
RA	Qy 61 ENYYWAASECMQDFNTMFTNCYIYNKPTDDIVMACTLEKIFLQVAVSMP 111
RA	Db 88 ENYYWAASECMQDFNTMFTNCYIYNKPTDDIVMAQALEKIFLQVAVSMP 138

Q7SX66	PRELIMINARY;	PRT;	558 AA.					
AC	Q7SX66;							
DT	01-OCT-2003	(TREMBREL 25, Created)						
DT	01-OCT-2003	(TREMBREL 25, Last sequence update)						
DT	01-OCT-2003	(TREMBREL 25, Last annotation update)						
DE	Hypothetical protein (Fragment).							
DB	Brachydanio rerio (zebrafish) (Danio rerio).							
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio; Danio.							
NCBI_TaxID	7855;							
OX	NCBI_TaxID:7855;							
RN	SEQUENCE FROM N.A.							
RP	RC							
STRAIN=AB; TISSUE=Body;								
MDLINE=22388257;								
RX	Medline=12477932;							
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shanmen C.M., Schuler G.D., Altschul S.F., Zeeberg B.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marszina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., McDonald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullally S.J., Bosak S.A., McEwan P.J., McEwan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Heilton E., Kettman J.M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young P.C., Shverchenko Y., Bouffard G.G., Blakesley R.W., Tschichman J.W., Green E.D., Dickson N.C., Rodriguez A.C., Grimes J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Grzimek J., Smaliski D.E., Schnierch A., Schein J.B., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).							
RN	[2]							
SEQUENCE FROM N.A.								
RP	SEQUENCE FROM N.A.							
RC	RA							
STRASBERG R.								
DT	Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.							
DR	EMBL; BC055508; AA055508.1; -							
KW	Hypothetical protein.							
PT	NON-TER							
SQ	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;					
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91;	Conservative 11;						
NON-TER	558	558						
SQ	SEQUENCE	558 AA;	62059 MW;	61681C83FB88FA3D CRC64;				
Query Match	85.5%	Score 517;	DB 13;	Length 558;				
Best Local Similarity	82.0%	Pred. No. 3.8e-49;	Mismatches 9;	Indels 0;	Gaps 0;			
Matches	91							

OM Protein - protein search, using sw model						
Run on: July 9, 2004, 13:05:11 ; Search time 23 Seconds (without alignments) 249.151 Million cell updates/sec						
Title:	US-09-784-553C-19					
Perfect score:	605					
Sequence:	1 KPGFVTNQLQYLHKVVMKAL,.....IVLMAQTLERIKFLQKVASMMP 111					
Scoring table:	BLOSUM62					
Gapop:	10.0 , Gapext 0.5					
Searched:	389414 seqs, 51625971 residues					
Total number of hits satisfying chosen parameters:	389414					
Minimum DB seq length: 0	Maximum DB seq length: 2000000000					
Post-processing: Minimum Match 0%	Maximum Match 100%					
	Listing first 45 summaries					
Database :						
	Issued Patents AA:*					
1	1: /cgn2_6_ptodata/2/iaa/5A_COMB.pep:*					
2	2: /cgn2_6_ptodata/2/iaa/5B_COMB.pep:*					
3	3: /cgn2_6_ptodata/2/iaa/6A_COMB.pep:*					
4	4: /cgn2_6_ptodata/2/iaa/6B_COMB.pep:*					
5	5: /cgn2_6_ptodata/2/iaa/207US_COMB.pep:*					
6	6: /cgn2_6_ptodata/2/iaa/backfiles1.pep:*					
	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES						
Result No.	Score	Query	Match Length	DB ID	Description	
1	605	100.0	754	4	US-09-392-714-20	Sequence 20, APP
2	542	89.6	726	4	US-09-392-714-21	Sequence 21, APP
3	500	82.6	722	4	US-09-392-714-22	Sequence 22, APP
4	459	75.9	947	4	US-09-418-780A-1	Sequence 1, APP
5	459	75.9	947	4	US-09-392-714-23	Sequence 23, APP
6	369	61.0	65	1	US-08-027-516-5	Sequence 5, APP
7	369	61.0	65	5	PCT-US55-04682-5	Sequence 6, APP
8	204	33.7	65	1	US-08-227-336-6	Sequence 6, APP
9	204	33.7	65	8	PCT-US55-04682-6	Sequence 6, APP
10	192	31.1	238	4	US-09-257-119-80	Sequence 80, APP
11	184.5	30.5	2441	1	US-08-194-668-2	Sequence 2, APP
12	184.5	30.5	2441	3	US-08-361-739-2	Sequence 2, APP
13	184.5	30.5	2441	4	US-09-514-241A-8	Sequence 2, APP
14	184.5	30.5	2441	4	US-09-086-316-2	Sequence 2, APP
15	181.5	30.0	2414	1	US-08-227-536-2	Sequence 2, APP
16	181.5	30.0	2414	5	PCT-US55-04682-2	Sequence 10, APP
17	180.5	29.8	2442	4	US-09-14-241A-10	Sequence 55, APP
18	158	26.1	61	4	US-09-418-710-55	Sequence 42, APP
19	153	25.3	1972	4	US-09-418-710-21	Sequence 21, APP
20	149	24.6	1972	4	US-09-418-710-21	Sequence 53, APP
21	142	23.5	559	4	US-09-418-710-53	Sequence 67, APP
22	139.5	23.1	1525	4	US-09-18-710-69	Sequence 27, APP
23	139.5	23.1	1527	4	US-09-418-710-27	Sequence 29, APP
24	139.5	23.1	1531	4	US-09-418-710-29	Sequence 72, APP
25	139	23.0	1693	4	US-09-418-710-72	Sequence 70, APP
26	22.6	1673	4	US-09-418-710-70	Sequence 1, APP	
27	16.5	1674	4	US-09-418-710-71	Sequence 1, APP	

RESULT 1  
US-09-332-714-20  
; Sequence 20 , Application US/09392714A  
; Patent No. 6666147  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; Gure, Ali O.  
; APPLICANT: Williamson, Barbara  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Cancer Associated Antigens and Uses  
; FILE REFERENCE: L0461/7062  
; CURRENT APPLICATION NUMBER: US/09/392,714A  
; CURRENT FILING DATE: 1999-09-09  
; EARLIER APPLICATION NUMBER: PCT/US99/14679  
; EARLIER FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 754  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-332-714-20

ALIGNMENTS

Sequence 14, App1  
Sequence 14, App1  
Sequence 14, App1  
Sequence 11, App1  
Sequence 13, App1  
Sequence 4, App1  
Sequence 4, App1  
Sequence 94, App1  
Sequence 71, App1  
Sequence 38, App1  
Sequence 50, App1  
Sequence 2, App1  
Sequence 51, App1  
Sequence 39, App1  
Sequence 52, App1  
Sequence 8, App1  
Sequence 8, App1

Query Match 100.0% ; Score 605; DB 4; Length 754;  
Best Local Similarity 100.0% ; Pred. No. 6.5e-71; Mismatches 0; Indels 0; Gaps 0

Query 1 KPGRTYNQQLHLKTYMKALWKHQAWPFPQDVAKLGLPDYHKLKIOPNDGMGTTKRL 60  
Db 24 KPGRTYNQQLHLKTYMKALWKHQAWPFPQDVAKLGLPDYHKLKIOPNDGMGTTKRL 83

Query 61 ENYYWAASECMDFTNTMFNTCYINKPTDDIVLMAQTLKEFLQVYASMP 111  
Db 84 ENYYWAASECMDFTNTMFNTCYINKPTDDIVLMAQTLKEFLQVYASMP 134

RESULT 2  
US-09-332-714-21  
; Sequence 21 , Application US/09392714A  
; Patent No. 66886147  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; Gure, Ali O.  
; APPLICANT: Williamson, Barbara  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Cancer Associated Antigens and Uses  
; TITLE OF INVENTION: Therefor

```

FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 21
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-21

Query Match 89.6%; Score 542; DB 4; Length 726;
Best Local Similarity 87.4%; Pred. No. 1.2e-62;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 KPGRVTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 60
Db 31 KPGRKTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 90
Qy 61 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMACTLEKIFLQVVASMP 111
Db 91 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMAQALEKIFLQVQASMP 141

RESULT 3
US-09-392-714-22
; Sequence 22, Application US/09392714A
; Patent No. 6386147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 22
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-22

Query Match 82.6%; Score 500; DB 4; Length 722;
Best Local Similarity 80.2%; Pred. No. 4.2e-57;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
Qy 1 KPGRVTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 60
Db 55 KPGRKTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 114
Qy 61 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMACTLEKIFLQVVASMP 111
Db 115 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMAQALEKIFLQVQASMP 165

RESULT 4
US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;
Qy 1 KPGRVTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 60
Db 24 KNGRLTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 83
Qy 61 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMACTLEKIFLQVVASMP 111
Db 84 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMAQALEKIFLQVQASMP 134

RESULT 5
US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 6386147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;
Qy 1 KPGRVTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 60
Db 24 KNGRLTNQOLYIHKVYMAWKHQFQAWPFPQVDALKLGPYKIKIKOPMDGTTIKRRL 83
Qy 61 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMACTLEKIFLQVVASMP 111
Db 84 ENYYWAASECMDNTMFNTCYIYNKPTDDIVLMAQALEKIFLQVQASMP 134

RESULT 6
US-08-227-536-5
; Sequence 5, Application US/08227536
; Patent No. 5658384
; GENERAL INFORMATION:
; APPLICANT: Beckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION

```

TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
 NUMBER OF SEQUENCES: 13  
 ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
 STREET: Ten Post Office Square  
 CITY: Boston  
 STATE: MA  
 ZIP: 02109  
 COUNTRY: US  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/227,536  
 FILING DATE: 14-APR-1994  
 CLASSIFICATION: 436  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Williams Ph.D., Kathleen A.  
 REGISTRATION NUMBER: 34,380  
 REFERENCE/DOCKET NUMBER: DFCI-308XXX  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-2280  
 TELEFAX: (617) 451-0313  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 65 amino acids  
 TYPE: amino acid  
 TOPOLogy: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 US-08-227-536-5

## RESULT 7

Query Match 61.0%; Score 369; DB 1; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 24 QFAWFRQPVDAVKGLPDYHKIKQPMDMGTTIKRLENNYWAASSCMQDNTMFNCY 83  
 Db 1 QFAWFRQPVDAVKGLPDYHKIKQPMDMGTTIKRLENNYWAASSCMQDNTMFNCY 60  
 Qy 84 IYNKP 88  
 Db 61 IYNKP 65

RESULT 8

Query Match 61.0%; Score 369; DB 5; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 QFAWFRQPVDAVKGLPDYHKIKQPMDMGTTIKRLENNYWAASSCMQDNTMFNCY 83  
 Db 1 QFAWFRQPVDAVKGLPDYHKIKQPMDMGTTIKRLENNYWAASSCMQDNTMFNCY 60

US-08-227-536-6

; Sequence 6, Application US/08227536  
 ; Patent No. 5658734  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eckner, Richard  
 ; APPLICANT: Ewen, Mark  
 ; APPLICANT: Livingston, David  
 ; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
 ; TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
 ; NUMBER OF SEQUENCES: 13  
 ; CURRENT APPLICATION DATA:  
 ; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
 ; STREET: Ten Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: US  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/227,536  
 ; FILING DATE: 14-APR-1994  
 ; CLASSIFICATION: 436  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Ph.D., Kathleen A.  
 ; REGISTRATION NUMBER: 34,380  
 ; REFERENCE/DOCKET NUMBER: DFCI-308XX  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-2290  
 ; TELEFAX: (617) 451-0313  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 65 amino acids  
 ; TYPE: amino acid

TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: internal  
 US-08-227-536-6

Query Match 33.7%; Score 204; DB 1; Length 65;  
 Best Local Similarity 57.8%; Pred. No. 1.8e-19;  
 Matches 37; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

RESULT 10  
 US-09-257-179-80  
 ; Sequence 80, Application US/09257179  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; PATENT NO.: 6410709  
 ; TITLE OF INVENTION: Human Secreted Proteins  
 ; FILE REFERENCE: P20015PL  
 ; CURRENT APPLICATION NUMBER: US/09/257,179  
 ; EARLIER APPLICATION NUMBER: PCT/US99/17709  
 ; EARLIER FILING DATE: 1999-02-25  
 ; EARLIER APPLICATION NUMBER: PCT/US99/17709  
 ; EARLIER FILING DATE: 1998-08-27  
 ; EARLIER APPLICATION NUMBER: 60/056,270  
 ; EARLIER FILING DATE: 1997-08-29  
 ; EARLIER APPLICATION NUMBER: 60/056,271  
 ; EARLIER FILING DATE: 1997-08-29  
 ; EARLIER APPLICATION NUMBER: 60/056,247  
 ; EARLIER FILING DATE: 1997-08-29  
 ; EARLIER APPLICATION NUMBER: 60/056,247  
 ; EARLIER FILING DATE: 1997-08-29  
 ; EARLIER APPLICATION NUMBER: 60/056,073  
 ; NUMBER OF SEQ ID NOS: 128  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 80  
 ; LENGTH: 238  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-257-179-80

RESULT 9  
 PCT-US95-04682-6  
 ; Sequence 6, Application PC/TUS9504682  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION  
 ; TITLE OF INVENTION: FACTOR B300 AND USES OF P300  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Weinhardt, Schurigin, Gagnebin & Hayes  
 ; STREET: Ten Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: US  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04682  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/227,536  
 ; FILING DATE: 14-April-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Holliday C. Heine, Ph.D.  
 ; REGISTRATION NUMBER: 34,346  
 ; REFERENCE/DOCKET NUMBER: DFCT-308Xq999  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 451-0313  
 ; TELEFAX: (617) 451-2290  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 65 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE: internal  
 ; PCT-US95-04682-6

Query Match 33.7%; Score 204; DB 5; Length 65;  
 Best Local Similarity 57.8%; Pred. No. 1.8e-19;  
 Matches 37; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

RESULT 11  
 US-08-194-468-2  
 ; Sequence 2, Application US/08194468  
 ; Patent No. 5750316  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Montmann, Marc R.  
 ; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF  
 ; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN  
 ; TITLE OF INVENTION: RESPONSIVE GENES  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
 ; STREET: 444 South Flower Street, Suite 2000  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/194,468  
 ; FILING DATE: 10-FEB-1994  
 ; CLASSIFICATION: 435

Query Match 33.7%; Score 204; DB 5; Length 65;  
 Best Local Similarity 57.8%; Pred. No. 1.8e-19;  
 Matches 37; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

RESULT 12  
 US-08-194-468-2  
 ; Sequence 2, Application US/08194468  
 ; Patent No. 5750316  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Montmann, Marc R.  
 ; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF  
 ; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN  
 ; TITLE OF INVENTION: RESPONSIVE GENES  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
 ; STREET: 444 South Flower Street, Suite 2000  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/194,468  
 ; FILING DATE: 10-FEB-1994  
 ; CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.  
 REGISTRATION NUMBER: 31,192  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) -546-4737  
 TELEFAX: (619) -546-9392  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2441 amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-194-468-2

Query Match 30.5%; Score 184.5; DB 1; Length 2441;  
 Best Local Similarity 38.0%; Pred. No. 7.8e-15; Indels 1; Gaps 1;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;  
 Qy 17 MKAIIWHQ-FAWPFRQPVDAVKLGLPDYHKKIKQPMGDKTIRRLENNYYWAASECMQDF 75  
 Db 1099 LEALYRQDPESLPFRQPVDPQQLGIPDYFIVKNPWDLSTIKRKLDGTQYQEPWQYDDV 1158  
 Qy 76 NTMFTNCYIYNKPTDDIVMAQTLLEKIFLQKV 107  
 Db 1159 RLMFNNAWLYNRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 12 US-08-961-739-2

Sequence 2, Application US/08961739A  
 Patent No. 6063583

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.  
 TITLE OF INVENTION: Methods for Treating Diabetes Mellitus  
 FILE REFERENCE: SALK1650-1  
 CURRENT APPLICATION NUMBER: US/08/961,739A  
 EARLIER APPLICATION NUMBER: US 194,468  
 EARLIER FILING DATE: 1994-02-10  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 2441  
 TYPE: PRT  
 ORGANISM: Mus  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1) .. (2441)  
 OTHER INFORMATION: Xaa = Any Amino Acid

US-08-961-739-2

Query Match 30.5%; Score 184.5; DB 3; Length 2441;  
 Best Local Similarity 38.0%; Pred. No. 7.8e-15; Indels 1; Gaps 1;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;  
 Qy 17 MKAIIWHQ-FAWPFRQPVDAVKLGLPDYHKKIKQPMGDKTIRRLENNYYWAASECMQDF 75  
 Db 1099 LEALYRQDPESLPFRQPVDPQQLGIPDYFIVKNPWDLSTIKRKLDGTQYQEPWQYDDV 1158  
 Qy 76 NTMFTNCYIYNKPTDDIVMAQTLLEKIFLQKV 107  
 Db 1159 RLMFNNAWLYNRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 13 US-09-514-247A-8

Sequence 8, Application US/09514247A  
 Patent No. 6365361  
 GENERAL INFORMATION:  
 APPLICANT: TANABE SEIYAKU CO. LTD.  
 APPLICANT: TANIGUCHI, Tencoyasu  
 APPLICANT: MIZUAKI, Junko

RESULT 15 US-08-227-536-2  
 Sequence 2, Application US/08227536  
 Patent No. 565784  
 GENERAL INFORMATION:

FILE REFERENCE: TANIGUCHI 6 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPAF ;  
 CURRENT APPLICATION NUMBER: US/09/514,247A ; CURRENT FILING DATE: 2000-02-28 ;  
 PRIORITY NUMBER: PCT/JP98/03734 ;  
 PRIOR FILING DATE: 1998-08-24 ;  
 PRIORITY NUMBER: JP231084/1997 ;  
 PRIOR FILING DATE: 1997-08-27 ;  
 PRIORITY NUMBER: JP NOS: 10 ;  
 SOFTWARE: Patentin version 3.0 ;  
 SEQ ID NO 8 ;  
 LENGTH: 2441 ;  
 TYPE: PRT ;  
 ORGANISM: mouse ;  
 US-09-514-247A-8

Query Match 30.5%; Score 184.5; DB 4; Length 2441;  
 Best Local Similarity 38.0%; Pred. No. 7.8e-15; Indels 1; Gaps 1;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;  
 Qy 17 MKAIIWHQ-FAWPFRQPVDAVKLGLPDYHKKIKQPMGDKTIRRLENNYYWAASECMQDF 75  
 Db 1099 LEALYRQDPESLPFRQPVDPQQLGIPDYFIVKNPWDLSTIKRKLDGTQYQEPWQYDDV 1158  
 Qy 76 NTMFTNCYIYNKPTDDIVMAQTLLEKIFLQKV 107  
 Db 1159 RLMFNNAWLYNRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 14 US-09-686-316-2  
 Sequence 2, Application US/09686316  
 Patent No. 6646115  
 GENERAL INFORMATION:  
 APPLICANT: Montminy, Marc R.  
 TITLE OF INVENTION: Methods for Treating Diabetes Mellitus  
 FILE REFERENCE: SALK1650-1  
 CURRENT APPLICATION NUMBER: US/09/686,316  
 CURRENT FILING DATE: 2000-10-10  
 PRIORITY NUMBER: US/08/961,739  
 PRIOR FILING DATE: 1997-10-31  
 PRIORITY NUMBER: US 194,468  
 PRIOR FILING DATE: 1994-01-10  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 2441  
 TYPE: PRT  
 ORGANISM: Mus  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1) .. (2441)  
 OTHER INFORMATION: Xaa = Any Amino Acid

US-09-686-316-2

Query Match 30.5%; Score 184.5; DB 4; Length 2441;  
 Best Local Similarity 38.0%; Pred. No. 7.8e-15; Indels 1; Gaps 1;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;  
 Qy 17 MKAIIWHQ-FAWPFRQPVDAVKLGLPDYHKKIKQPMGDKTIRRLENNYYWAASECMQDF 75  
 Db 1099 LEALYRQDPESLPFRQPVDPQQLGIPDYFIVKNPWDLSTIKRKLDGTQYQEPWQYDDV 1158  
 Qy 76 NTMFTNCYIYNKPTDDIVMAQTLLEKIFLQKV 107  
 Db 1159 RLMFNNAWLYNRKTSRVYKFCSKLAEVFEQEI 1190

APPLICANT: Eckner, Richard  
 APPLICANT: Ewen, Mark  
 APPLICANT: Livingston, David  
 TITLE OF INVENTION: NUCLEAR ACID, ENCODING TRANSCRIPTION  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes  
 STREET: Ten Post Office Square  
 CITY: Boston  
 STATE: MA  
 COUNTRY: US  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/227,536  
 FILING DATE: 14-APR-1994  
 CLASSIFICATION: 436  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Williams Ph.D., Kathleen A.  
 REGISTRATION NUMBER: 34,380  
 REFERENCE/DOCKET NUMBER: DECI-308XX  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-2290  
 TELEFAX: (617) 451-0313  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2414 amino acids  
 TYPE: amino acid  
 TOPOLOGY: Linear  
 MOLECULE TYPE: protein  
 US-08-227-536-2

Query Match 30.0% Score 181.5; DB 1; Length 2414;  
 Best Local Similarity 38.0%; Pred. No. 1.9e-14; Gaps 1;  
 Matches 35; Conservative 20; Mismatches 36; Indels 1;  
 Qy 17 MKALWKGQ-FAWPPFQPVDAVKLGPDPYHKIKOPMDNGTIKKRLENNYWAASECMQDF 75  
 Db 1062 LEALYRQDPPLPFFQPVDDQQLGIPDVKSPMDLSTIKRKLDTQYQSPWQYVDDI 1121  
 Qy 76 NTMFNCYIYNKPFDIVLMAQTEKIFQKV 107  
 Db 1122 WLMFNRNAWLYNRKTSRVYKCSKLESEVPEQEI 1153

Search completed: July 9, 2004, 13:08:43  
 Job time : 24 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
 Run on: July 9, 2004, 13:07:41 ; Search time 47 Seconds  
 (without alignments)

736,658 Million cell updates/sec

Title: US-09-784-553C-19  
 Perfect score: 605  
 Sequence: 1 K2RVTNOLQYLHKVVMKAL.....IVIMAQTLKIFLQKVASMP 111

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311910243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0 %

Maximum Match 100 %

Listing First 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpa/us07\_pubcomb.pep:\*

2: /cgn2\_6/ptodata/1/pubpa/pcr\_new\_pub.pep:\*

3: /cgn2\_6/ptodata/1/pubpa/us06\_new\_pub.pep:\*

4: /cgn2\_6/ptodata/1/pubpa/us06\_pubcomb.pep:\*

5: /cgn2\_6/ptodata/1/pubpa/pcr5\_pubcomb.pep:\*

6: /cgn2\_6/ptodata/1/pubpa/pcr5\_pubcomb.pep:\*

7: /cgn2\_6/ptodata/1/pubpa/us08\_new\_pub.pep:\*

8: /cgn2\_6/ptodata/1/pubpa/us09\_pubcomb.pep:\*

9: /cgn2\_6/ptodata/1/pubpa/us09b\_pubcomb.pep:\*

10: /cgn2\_6/ptodata/1/pubaa/us09c\_pubcomb.pep:\*

11: /cgn2\_6/ptodata/1/pubaa/us09d\_pubcomb.pep:\*

12: /cgn2\_6/ptodata/1/pubaa/us09e\_pubcomb.pep:\*

13: /cgn2\_6/ptodata/1/pubaa/us10a\_pubcomb.pep:\*

14: /cgn2\_6/ptodata/1/pubaa/us10b\_pubcomb.pep:\*

15: /cgn2\_6/ptodata/1/pubaa/us10c\_pubcomb.pep:\*

16: /cgn2\_6/ptodata/1/pubaa/us10d\_pubcomb.pep:\*

17: /cgn2\_6/ptodata/1/pubaa/us60\_new\_pub.pep:\*

18: /cgn2\_6/ptodata/1/pubaa/us60\_pubcomb.pep:\*

Prd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	111	12 US-09-784-553C-19	Sequence 19, Appl
2	605	100.0	111	15 US-10-209-201C-19	Sequence 19, Appl
3	605	100.0	801	14 US-10-145-473-12	Sequence 42, Appl
4	542	89.6	111	12 US-09-784-553C-20	Sequence 20, Appl
5	542	89.6	111	15 US-10-209-201C-0	Sequence 20, Appl
6	542	89.6	140	9 US-09-764-864-1772	Sequence 1572, Ap
7	542	89.6	235	9 US-09-764-864-1159	Sequence 1159, Ap
8	500	82.5	1390	12 US-10-092-900A-224	Sequence 224, Appl
9	498	82.3	111	12 US-09-784-553C-21	Sequence 21, Appl
10	498	82.3	111	15 US-10-209-201C-21	Sequence 21, Appl
11	459	75.9	947	14 US-10-293-822-1	Sequence 1, Appl
12	269	44.5	113	12 US-09-784-553C-23	Sequence 23, Appl
13	269	44.5	113	15 US-10-209-201C-23	Sequence 25, Appl
14	256.5	42.4	113	15 US-09-784-553C-25	Sequence 25, Appl
15	256.5	42.4	113	15 US-10-209-201C-25	Sequence 25, Appl

RESULT 1  
 US-09-784-553C-19  
 ; Sequence 19, Application US-09784553C  
 ; Publication No. US20040043378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZHOU, MING-MING  
 ; AGGAWAL, ANEEL  
 ; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS  
 ; FILE REFERENCE: 2459-1-003 CIP  
 ; CURRENT APPLICATION NUMBER: US-09-784-553C-2  
 ; PRIORITY APPLICATION NUMBER: 09/510,314  
 ; PRIORITY FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 63  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 19  
 ; LENGTH: 111  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-09-784-553C-19

Query Match	Match	Score	605;	DB 12;	Length	111;
Best	Local Similarity	100.0 %	Pred. No. 5.6-66;			
Matches	111;	Conservative	0;	Missmatches	0;	Gaps 0;
Qy	1	KPGRVNTQLOVHLKVKMALKHOFAMPFROPVDAVTKLGLIPDYHKLIKQPMDMGTIKRRL 60				
Db	1	KPGRVNTQLOVHLKVKMALKHOFAMPFROPVDAVTKLGLIPDYHKLIKQPMDMGTIKRRL 60				
Qy	61	ENYYWAASECMDNTMFTNCKYIYAKPTDDIVLNQTLKIFLQKVASMP 111				
Db	61	ENYYWAASECMDNTMFTNCKYIYAKPTDDIVLNQTLKIFLQKVASMP 111				

RESULT 2  
 US-10-209-201C-19  
 ; Sequence 19, Application US-10209201C  
 ; Publication No. US2004009613A1

GENERAL INFORMATION:  
 ; APPLICANT: Verdin, Eric  
 ; APPLICANT: Bruland, Joan  
 ; APPLICANT: Ott, Melanie  
 ; APPLICANT: Zhou, Ming-Ming  
 ; APPLICANT: Aggarwal, Aneel  
 ; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains  
 ; FILE REFERENCE: 2459-1-003CIPDIV  
 ; CURRENT APPLICATION NUMBER: US710/209, 2011  
 ; PRIOR APPLICATION NUMBER: 09/784, 553  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: 09/510, 314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 19  
 ; LENGTH: 111  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-209-201-C-19

Query Match 100.0%; Score 605; DB 15; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-66;  
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KPGRVTNQQLHKKVVMKALWKHQFAWPFFQPDAVKLGLPDYHKIJKOPMDNGTIRKL 60  
 Db 1 KPGRVTNQQLHKKVVMKALWKHQFAWPFFQPDAVKLGLPDYHKIJKOPMDNGTIRKL 60  
 Qy 61 ENNYWAASECMDENTMFNCYIYNKPTDDIVLMAQTLKEFLQKVASMP 111  
 Db 61 ENNYWAASECMDENTMFNCYIYNKPTDDIVLMAQTLKEFLQKVASMP 111

RESULT 3  
 US-10-146-473-42

; Sequence 42, Application US/10146473

; Publication No. US20030108888A1

GENERAL INFORMATION:  
 ; APPLICANT: Scianian, Matthew

; APPLICANT: Gout, Ivan

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Gure, Al

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Old, Lloyd

; TITLE OF INVENTION: Breast Cancer Antigens

; FILE REFERENCE: L00441/70130 (JRV)

; CURRENT APPLICATION NUMBER: US10/146,473

; PRIOR APPLICATION NUMBER: US 2002-05-15

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 82

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 42

; LENGTH: 801

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match 100.0%; Score 605; DB 14; Length 801;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-65;  
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KPGRVTNQQLHKKVVMKALWKHQFAWPFFQPDAVKLGLPDYHKIJKOPMDNGTIRKL 60  
 Db 71 KPGRVTNQQLHKKVVMKALWKHQFAWPFFQPDAVKLGLPDYHKIJKOPMDNGTIRKL 130  
 Qy 61 ENNYWAASECMDENTMFNCYIYNKPTDDIVLMAQTLKEFLQKVASMP 111  
 Db 131 ENNYWAASECMDENTMFNCYIYNKPTDDIVLMAQTLKEFLQKVASMP 181

RESULT 4  
 US-09-784-553C-20

; Sequence 20, Application US/09784553C

; Publication No. US20040043378A1

GENERAL INFORMATION:  
 ; APPLICANT: Zhou, Ming-Ming

; APPLICANT: Aggarwal, Aneel

; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains

; FILE REFERENCE: 2459-1-003CIPDIV

; CURRENT APPLICATION NUMBER: US710/209, 2011

; PRIOR APPLICATION NUMBER: 09/784, 553

; PRIOR FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: 09/510, 314

; PRIOR FILING DATE: 2000-02-22

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 20

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-784-553C-20

; Sequence 20, Application US/09784553C

; Publication No. US20040043378A1

GENERAL INFORMATION:  
 ; APPLICANT: Verdin, Eric

; APPLICANT: Bruland, Joan

; APPLICANT: Ott, Melanie

; APPLICANT: Zhou, Ming-Ming

; APPLICANT: Aggarwal, Aneel

; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains

; FILE REFERENCE: 2459-1-003CIPDIV

; CURRENT APPLICATION NUMBER: US710/209, 2011

; PRIOR APPLICATION NUMBER: 09/784, 553

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: 09/510, 314

; PRIOR FILING DATE: 2000-02-22

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 20

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-784-553C-20

; Sequence 20, Application US/09784553C

; Publication No. US20040043378A1

GENERAL INFORMATION:  
 ; APPLICANT: Zhou, Ming-Ming

; APPLICANT: Aggarwal, Aneel

; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains

; FILE REFERENCE: 2459-1-003CIPDIV

; CURRENT APPLICATION NUMBER: US710/209, 2011

; PRIOR APPLICATION NUMBER: 09/784, 553

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 20

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Homo sapiens

RESULT 5  
 US-10-209-201C-20

; Sequence 20, Application US/10209201C

; Publication No. US2004009613A1

GENERAL INFORMATION:  
 ; APPLICANT: Verdin, Eric

; APPLICANT: Bruland, Joan

; APPLICANT: Ott, Melanie

; APPLICANT: Zhou, Ming-Ming

; APPLICANT: Aggarwal, Aneel

; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains

; FILE REFERENCE: 2459-1-003CIPDIV

; CURRENT APPLICATION NUMBER: US710/209, 2011

; PRIOR APPLICATION NUMBER: 09/784, 553

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: 09/510, 314

; PRIOR FILING DATE: 2000-02-22

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 20

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-209-201C-20

; Sequence 20, Application US/10209201C

; Publication No. US2004009613A1

GENERAL INFORMATION:  
 ; APPLICANT: Zhou, Ming-Ming

; APPLICANT: Aggarwal, Aneel

; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains

; FILE REFERENCE: 2459-1-003CIPDIV

; CURRENT APPLICATION NUMBER: US710/209, 2011

; PRIOR APPLICATION NUMBER: 09/784, 553

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 20

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Homo sapiens

RESULT 6  
 US-09-764-866-1572  
 Sequence 1572, Application US/09764864  
 Patent No. US200112753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT223  
 CURRENT APPLICATION NUMBER: US/09/764, 864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1152  
 LENGTH: 140  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-764-866-1572

Query Match 89.6%; Score 542; DB 9; Length 140;  
 Best Local Similarity 87.4%; Pred. No. 3.9e-58;  
 Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
 Qy 1 KPGRTNQLQQLHVKWQPAWPRQDAVKGQLPDYHKKIKQPMGTTKRL 60  
 Db 5 KPGRTNQLQQLHVKWQPAWPRQDAVKGQLPDYHKKIKQPMGTTKRL 64  
 Qy 67 ENNYWAAASBCMODNTMTCYLYNKPTDDIVLMAQLEKIFLQKVASMP 111  
 Db 65 ENNYWAAASBCMODNTMTCYLYNKPTDDIVLMAQLEKIFLQKVASMP 115

RESULT 7  
 US-09-764-866-1159  
 Sequence 1159, Application US/09764864  
 Patent No. US200112753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT223  
 CURRENT APPLICATION NUMBER: US/09/764, 864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1159  
 LENGTH: 235  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: SITE  
 NAME/KEY: SITE  
 LOCATION: (215)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (221)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-764-864-1159

Query Match 89.6%; Score 542; DB 9; Length 235;

Best Local Similarity 87.4%; Pred. No. 7.3e-58;

Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

RESULT 8  
 US-10-092-900A-224  
 Sequence 224, Application US/10092900A  
 Publication No. US2004004382A1  
 GENERAL INFORMATION:  
 APPLICANT: Padigaru, Muralidhara  
 APPLICANT: Spytek, Kimberly A.  
 APPLICANT: Shenvoy, Suresh G.  
 APPLICANT: Taupier Jr., Raymond J.  
 APPLICANT: Pena, Carol E.A.  
 APPLICANT: Li, Li  
 APPLICANT: Zerhusen, Bryant D.  
 APPLICANT: Gusev, Vladimir Y.  
 APPLICANT: Ji, Weizhen  
 APPLICANT: Gorman, Linda  
 APPLICANT: Miller, Charles E.  
 APPLICANT: Kekuda, Ramesh  
 APPLICANT: Patturajan, Meera  
 APPLICANT: Gangolli, Esha A.  
 APPLICANT: Vernet, Corine A.M.  
 APPLICANT: Guo, Xiaotia Sasha  
 APPLICANT: Tchernov, Velizar T.  
 APPLICANT: Fernandes, Elma R.  
 APPLICANT: Casman, Stacie J.  
 APPLICANT: Malanjkar, Uriel M.  
 APPLICANT: Gerach, Valerie  
 APPLICANT: Liu, Yi  
 APPLICANT: Anderson, David W.  
 APPLICANT: Spaderna, Steven K.  
 APPLICANT: Catterton, Elma  
 APPLICANT: Leite, Mario W.  
 APPLICANT: Zhong, Haihong  
 APPLICANT: Alsobrook, John P.  
 APPLICANT: Lepley, Denise M.  
 APPLICANT: Rieger, Daniel K.  
 APPLICANT: Burgess, Catherine E.  
 TITLE OF INVENTION: No. US2004004382A1 Proteins and Nucleic Acids Encoding Same  
 FILE REFERENCE: 21402-290C  
 CURRENT APPLICATION NUMBER: US/10/092, 900A  
 CURRENT FILING DATE: 2002-03-07  
 PRIOR APPLICATION NUMBER: USN 60/274, 322  
 PRIOR FILING DATE: 2001-03-08  
 PRIOR APPLICATION NUMBER: USN 60/283, 675  
 PRIOR FILING DATE: 2001-04-13  
 PRIOR APPLICATION NUMBER: USN 60/338, 092  
 PRIOR FILING DATE: 2001-12-03  
 PRIOR APPLICATION NUMBER: USN 60/274, 281  
 PRIOR FILING DATE: 2001-03-08  
 PRIOR APPLICATION NUMBER: USN 60/274, 191  
 PRIOR FILING DATE: 2001-03-08  
 PRIOR APPLICATION NUMBER: USN 60/325, 681  
 PRIOR FILING DATE: 2001-09-27  
 PRIOR APPLICATION NUMBER: USN 60/304, 354  
 PRIOR FILING DATE: 2001-07-10  
 PRIOR APPLICATION NUMBER: USN 60/279, 995  
 PRIOR FILING DATE: 2001-03-30  
 PRIOR APPLICATION NUMBER: USN 60/294, 899  
 PRIOR FILING DATE: 2001-05-31  
 PRIOR APPLICATION NUMBER: USN 60/287, 424  
 PRIOR FILING DATE: 2001-04-30  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SEQ ID NO 224  
 LENGTH: 1390  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-092-900A-224

Query Match 82.6%; Score 500; DB 12; Length 1390;

Best Local Similarity 80.2%; Pred. No. 8.9e-52;

Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KPGRTNQLQQLHVKWQPAWPRQDAVKGQLPDYHKKIKQPMGTTKRL 60  
 Db 5 KPGRTNQLQQLHVKWQPAWPRQDAVKGQLPDYHKKIKQPMGTTKRL 64  
 Qy 61 ENNYWAAASBCMODNTMTCYLYNKPTDDIVLMAQLEKIFLQKVASMP 111  
 Db 65 ENNYWAAASBCMODNTMTCYLYNKPTDDIVLMAQLEKIFLQKVASMP 115

RESULT 9  
 US-09-784-553C-21  
 ; Sequence 21, Application US/09784553C  
 ; Publication No. US20040043378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZHOU, MING-MING  
 ; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS  
 ; CURRENT APPLICATION NUMBER: US/09/784,553C  
 ; FILE REFERENCE: 2459-1-003 CIP  
 ; CURRENT FILING DATE: 2003-07-08  
 ; PRIORITY NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 63  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 111  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-784-553C-21  
 Query Match 82.3%; Score 498; DB 12; Length 111;  
 Best Local Similarity 77.5%; Pred. No. 7.1e-53;  
 Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;  
 Db 1 KPGRVNQLQYLHKVYMKALWKHQPAWPRQPVDAVKLGPDYHKIKIKOPMDGTTKRL 60  
 1 RGRNTNQLQYLHKVYMKALWKHQPAWPRQPVDAVKLNPWPPQPVDAKLNPKDYHKIKIKOPMDGTTKRL 60

Qy 61 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111  
 Db 61 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111

RESULT 10  
 US-10-209-201C-21  
 ; Sequence 21, Application US/10209201C  
 ; Publication No. US2004009613A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bruland, Joan  
 ; APPLICANT: Verdin, Eric  
 ; APPLICANT: Ott, Melanie  
 ; APPLICANT: Aggarwal, Annel  
 ; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains  
 ; CURRENT APPLICATION NUMBER: US/03/10209,201C  
 ; FILE REFERENCE: 2459-1-003 CIP  
 ; CURRENT FILING DATE: 2002-07-31  
 ; PRIORITY NUMBER: 09/714,553  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIORITY NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 111  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-209-201C-21  
 Query Match 82.3%; Score 498; DB 15; Length 111;  
 Best Local Similarity 77.5%; Pred. No. 7.1e-53;  
 Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KPGRVNQLQYLHKVYMKALWKHQPAWPRQPVDAVKLGPDYHKIKIKOPMDGTTKRL 60  
 1 RGRNTNQLQYLHKVYMKALWKHQPAWPRQPVDAVKLNPWPPQPVDAKLNPKDYHKIKIKOPMDGTTKRL 60  
 1 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111  
 Db 1 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111  
 1 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111  
 Db 84 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 114

RESULT 11  
 US-10-293-822-1  
 ; Sequence 1, Application US/10293822  
 ; Publication No. US20030083470A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jones, Michael H  
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR  
 ; FILE REFERENCE: 06/01-043001  
 ; CURRENT APPLICATION NUMBER: US/10/293,822  
 ; CURRENT FILING DATE: 2002-11-12  
 ; PRIORITY NUMBER: US/09/418,780  
 ; PRIOR APPLICATION NUMBER: 1999-10-15  
 ; PRIOR FILING DATE: 1999-10-15  
 ; PRIORITY NUMBER: PCT/JP98/01782  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIORITY NUMBER: JAPAN 9/116402  
 ; PRIOR FILING DATE: 1997-04-18  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 1  
 ; LENGTH: 947  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-293-822-1  
 Query Match 75.9%; Score 459; DB 14; Length 947;  
 Best Local Similarity 73.0%; Pred. No. 5.8e-47;  
 Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

Qy 1 KGRVNTQQLQYLHKVYMKALWKHQPAWPRQPVDAVKLGPDYHKIKIKOPMDGTTKRL 60  
 24 RGRNLTLQQLQYVVKLWKHQPAWPRQPVDAVKLGPDYHKIKIKOPMDGTTKRL 83  
 Qy 61 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 111  
 Db 84 ENNYWAASECQMDNTMFNTCNYKPTDVLMAQTLKFLQKVASMP 114

RESULT 12  
 US-09-784-553C-23  
 ; Sequence 23, Application US/09784553C  
 ; Publication No. US2004043378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZHOU, MING-MING  
 ; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS  
 ; FILE REFERENCE: 2459-1-003 CIP  
 ; CURRENT FILING DATE: 2003-07-08  
 ; PRIORITY NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 63  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 23  
 ; LENGTH: 113  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-784-553C-23  
 Query Match 44.5%; Score 269; DB 12; Length 113;  
 Best Local Similarity 46.0%; Pred. No. 7.8e-25;  
 Matches 52; Conservative 21; Mismatches 38; Indels 2; Gaps 1;

Qy 1 KGRVNTQQLQYLHKVYMKALWKHQPAWPRQPVDAVKLGPDYHKIKIKOPMDGTTKRL 58

Db 1 KGGKLSEQIKHNGILKELSLKHKRAYAWPFYKPVDAASALGHHDYHDIKKHMPDLSTVKR 60  
 Qy 59 RLENNYVPAASECMQDFNTMFTNCYIYNKPTDDIVMAQTLERKIFLQKVASMP 111  
 :  
 Db 61 KMENRDYRDAQEFAADYRLLMFSNCYKYNPPDHVVAMARKLQDVFEEFRYAKMP 113

RESULT 13  
 US-10-209-201C-23  
 ; Sequence 23, Application US/10209201C  
 ; Publication No. US20040009613A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Verdin, Eric  
 ; APPLICANT: Bruland, Joan  
 ; APPLICANT: Ott, Melanie  
 ; APPLICANT: Zhou, Ming-Ming  
 ; APPLICANT: Agarwall, Annel  
 ; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains  
 ; FILE REFERENCE: 2459-1-003CIPDIV  
 ; CURRENT APPLICATION NUMBER: US/10/209,201C  
 ; CURRENT FILING DATE: 2002-07-31  
 ; PRIOR APPLICATION NUMBER: 09/784,553  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 23  
 ; LENGTH: 113  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-209-201C-23

Query Match 44.5%; Score 269; DB 15; Length 113;  
 Best Local Similarity 46.0%; Pred. No. 7.8e-25; Indels 2; Gaps 1;  
 Matches 52; Conservative 21; Mismatches 38; Indels 2; Gaps 1;  
 Qy 1 KPGRTVNOQYLLHKVVMKALWKHQ -FAWPFROPVDAYKLGLPDYHKLIKQEMDMGTR 58  
 Db 1 KGGKLSEQIKHNGILKELSLKHKRAYAWPFYKPVDAASALGHHDYHDIKKHMPDLSTVKR 60

Qy 59 RLENNYVPAASECMQDFNTMFTNCYIYNKPTDDIVMAQTLERKIFLQKVASMP 111  
 Db 61 KMENRDYRDAQEFAADYRLLMFSNCYKYNPPDHVVAMARKLQDVFEEFRYAKMP 113

RESULT 14  
 US-09-784-553C-25  
 ; Sequence 25, Application US/09784553C  
 ; Publication No. US20040013378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZHOU, MING-MING  
 ; APPLICANT: AGARWALL, ANNEL  
 ; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS  
 ; FILE REFERENCE: 2459-1-003 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/784,553C  
 ; CURRENT FILING DATE: 2003-07-08  
 ; PRIOR APPLICATION NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 63  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 25  
 ; LENGTH: 113  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-784-553C-25

Query Match 42.4%; Score 256.5; DB 12; Length 113;  
 Best Local Similarity 48.5%; Pred. No. 2.7e-23;  
 Matches 48; Conservative 20; Mismatches 28; Indels 3; Gaps 2;  
 Qy 16 VMKALW--KHO-FAWPFROPVDAYKLGLPDYHKLIKQEMDMGTR 72  
 Db 15 ILKELFSKGEHSQYAMPFYKPVDAEMGLHDYHDIKKHMPDLGTVKR 74

Qy 73 QDFNTMFTNCYIYNKPTDDIVMAQTLERKIFLQKVASMP 111  
 Db 75 ADVLIFTNCYKINPPDHVVAMGRKLQDVFEEFRYANIP 113

Search completed: July 9, 2004, 13:13:32  
 Job time : 48 secs

Db 15 ILKELFSKGEHSQYAMPFYKPVDAEMGLHDYHDIKKHMPDLGTVKR 74  
 Qy 73 QDFNTMFTNCYIYNKPTDDIVMAQTLERKIFLQKVASMP 111  
 :  
 Db 75 ADVLIFTNCYKINPPDHVVAMGRKLQDVFEEFRYANIP 113

RESULT 15  
 US-10-209-201C-25  
 ; Sequence 25, Application US/10209201C  
 ; Publication No. US20040009613A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Verdin, Eric  
 ; APPLICANT: Bruland, Joan  
 ; APPLICANT: Ott, Melanie  
 ; APPLICANT: Zhou, Ming-Ming  
 ; APPLICANT: Agarwall, Annel  
 ; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains  
 ; FILE REFERENCE: 2459-1-003CIPDIV  
 ; CURRENT APPLICATION NUMBER: US/10/209,201C  
 ; CURRENT FILING DATE: 2002-07-31  
 ; PRIOR APPLICATION NUMBER: 09/784,553  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: 09/510,314  
 ; PRIOR FILING DATE: 2000-02-22  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 25  
 ; LENGTH: 113  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-209-201C-25

Query Match 42.4%; Score 256.5; DB 15; Length 113;  
 Best Local Similarity 48.5%; Pred. No. 2.7e-23;  
 Matches 48; Conservative 20; Mismatches 28; Indels 3; Gaps 2;  
 Qy 16 VMKALW--KHO-FAWPFROPVDAYKLGLPDYHKLIKQEMDMGTR 72  
 Db 15 ILKELFSKGEHSQYAMPFYKPVDAEMGLHDYHDIKKHMPDLGTVKR 74

Qy 73 QDFNTMFTNCYIYNKPTDDIVMAQTLERKIFLQKVASMP 111  
 Db 75 ADVLIFTNCYKINPPDHVVAMGRKLQDVFEEFRYANIP 113

